

Evaluation of the Presentation of Network Data via Visualization Tools for Network Analysts

by Renée E. Etoty, Dr. Robert F. Erbacher, and Dr. Christopher Garneau

ARL-TR-6865

March 2014

NOTICES

Disclaimers

The findings in this report are not to be construed as an official Department of the Army position unless so designated by other authorized documents.

Citation of manufacturer's or trade names does not constitute an official endorsement or approval of the use thereof.

Destroy this report when it is no longer needed. Do not return it to the originator.

Army Research Laboratory

Adelphi, MD 20783-1197

ARL-TR-6865

March 2014

Evaluation of the Presentation of Network Data via Visualization Tools for Network Analysts

Renée E. Etoty, Dr. Robert F. Erbacher, and Dr. Christopher Garneau
Computational and Information Sciences Directorate, ARL

REPORT DOCUMENTATION PAGE			Form Approved OMB No. 0704-0188		
<p>Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden, to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.</p> <p>PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.</p>					
1. REPORT DATE (DD-MM-YYYY) March 2014		2. REPORT TYPE		3. DATES COVERED (From - To) 09/2013 to 09/2014	
4. TITLE AND SUBTITLE Evaluation of the Presentation of Network Data via Visualization Tools for Network Analysts			5a. CONTRACT NUMBER		
			5b. GRANT NUMBER		
			5c. PROGRAM ELEMENT NUMBER		
6. AUTHOR(S) Renée E. Etoty, Dr. Robert F. Erbacher, and Dr. Christopher Garneau			5d. PROJECT NUMBER		
			5e. TASK NUMBER		
			5f. WORK UNIT NUMBER		
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) U.S. Army Research Laboratory ATTN: RDRL-CIN-D 2800 Powder Mill Road Adelphi, MD 20783-1197			8. PERFORMING ORGANIZATION REPORT NUMBER ARL-TR-6865		
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)			10. SPONSOR/MONITOR'S ACRONYM(S)		
			11. SPONSOR/MONITOR'S REPORT NUMBER(S)		
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution unlimited.					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT <p>In response to chaotic nature of network traffic, making it very difficult to differentiate normal from malicious traffic, we have designed a user study that tests the effectiveness and usefulness of tabular versus graphical displays on such data. The U.S. Army Research Laboratory's (ARL) in-house defense service providers are expert subjects, who undergo a simplified version of their computer network defense (CND) analyst tasks. We use their performance to acquire initial insights to their interpretation of display components, cognitive processes, and contextual knowledge. We quantitatively compare tabular versus graphical displays and compare their feedback with that of students, who serve as primary test subjects for developing visual displays for network monitoring. In this study, all participants act as analysts; their job is to identify evidence of compromise within a dataset of intrusion attempts on the fabricated network visually provided. We observe the participants' responses to the pattern matching activity created with interacting with the visual displays. The design variables are the distinct graphical layouts: tabular, parallel coordinates, and node-link. The response variables are true positive and false positive rates of event identification, the time required for event identification, and the qualitative questionnaire. Results help us better understand which of the visual layouts is most effective and useful for predicting cyber attacks.</p>					
15. SUBJECT TERMS Cyber Security, Cognitive Reasoning Process, Empirical Study, Information Visualization					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT UU	18. NUMBER OF PAGES 62	19a. NAME OF RESPONSIBLE PERSON Renée E. Etoty
a. REPORT Unclassified	b. ABSTRACT Unclassified	c. THIS PAGE Unclassified			19b. TELEPHONE NUMBER (Include area code) (301) 394-1835

Contents

List of Figures	iv
1. Introduction	1
2. Information Visualization	2
2.1 Visual Representation in Table Form.....	2
2.2 Visual Representation in Graphical Form.....	3
3. User Studies	4
3.1 Study Development	5
3.2 Experimental Design	8
4. Future Work	11
5. References	12
Appendix A. Approved Protocol Project#: ARL 13-050	15
Distribution List	56

List of Figures

Figure 1. A screenshot of the node-link graphical representation of computer network alerts. The user is asked here to determine regions of the visualization that imply intrusions (True Positive [TP]) and intrusion attempts (False Positives [FP]) by clicking near a particular link or node.	7
Figure 2. A screenshot of the tabular representation of computer network alerts. The user is asked here to determine which alert messages in the table imply intrusions (True Positive [TP]) and intrusion attempts (False Positives [FP]) by clicking the checkboxes in the ‘Suspicious’ column.	7
Figure 3. Organization chart of an overview of the development for this study.	8
Figure 4. Tabular Display, representation A.	9
Figure 5. Parallel Coordinates Display, representation B.	9
Figure 6. Node-link Display, representation C.	10
Figure 7. Experiment Design Overview.	11
Figure A-1. Tabular Display, representation A.	22
Figure A-2. Parallel Coordinates Display, representation B.	22
Figure A-3. Node-link Display, representation C.	22
Figure A-4. Alerts Example in a Tabular Display.	23
Figure A-5. (a) Line Visualizations Used for Parallel Coordinate Display. (b) Meanings of Line Visualizations Used (<i>II</i>).	24
Figure A-6. (a) Line Visualizations Used for Node-Link Display. (b) Meanings of Line Visualizations Used (<i>II</i>).	25
Figure A-7. (a) MATLAB shot of plot library being used. (b) MATHLAB highlights.	26
Figure A-8. (a) This is a screenshot of the “Introduction Overlay”. It introduces the subject to the mission, provides instructions on identifying threats, and highlights the features/functions of the game. (b) This is a screenshot of the “Exercise Overlay”. Here the subjects will be able to view, explore, and look deeper into the dataset via the particular visualization display. It is here after exploration that the subject determines and identifies the network threat. (c) This is a screenshot of the “Results Overlay”. A running log is kept in the background to keep track of each subject's performance. A module contains all three overlays and repeats three times for the three different visual displays. The subject's performance is displayed at the end of each module and once more at the end of the entire session for their overall time and accuracy performance.	28
Figure A-9. Experience-aided reasoning support system overview.	29

1. Introduction

The defense of computer networks incorporates network monitoring as a critical component for which the U.S. Army Research Laboratory (ARL) has become well-respected as result of its in-house computer network defense service provider. This network monitoring places heavy demands on the human analyst to identify and analyze threats, especially advanced persistent threats. Such threats require the analyst to correlate temporally and physically disparate events cognitively. The chaotic nature of network traffic data makes it very difficult to differentiate normal from malicious traffic.

Analysts have a difficult task characterized by the need to integrate technical knowledge with contextual knowledge under severe constraints. We plan to turn this cognitive overload into an opportunity by enhancing the visual displays used by analysts to create tools that more effectively reduce the cognitive load while directly aiding the correlation of data through visual organization of the data. Analysts typically work with tabular displays or raw data for conducting their tasks. Other types of displays providing more representations that are abstract may provide more insight into big data inter-relationships, patterns, and areas of interest.

The goal of this research is to examine and lay out the underlying science and theory of network-based intrusion detection—i.e., to develop a rigorous science of intrusion detection. Current techniques being developed in a very ad-hoc fashion have very little relevance to the real world or any real expectation that the developed techniques will be successful or useful. This can be seen in particular with visualization techniques. Numerous visualization techniques have been developed over the past decade, but we have not been able to identify any that have ever been successfully deployed. It is actually questionable whether analysts have even seen the majority of these techniques. Yet new techniques are consistently being designed and published. A further complication is the fact that developed techniques are not being tested with real-world data and will likely fail in the real world—i.e., with ARL Computer Network Defense Service Provider (CNDSP) data. We do not know why these techniques are failing to be deployed:

- Are analysts being given the opportunity to employ the techniques?
- Do the techniques meet analysts' needs or expectations?
- Do the techniques scale to the size of ARL data successfully?
- Are the techniques actually identifying relevant malicious activity?
- Will the techniques reduce analyst time or increase it?

In response, we have designed a user-study based on a cyber-security analysis game scenario and questionnaires to acquire initial insights from real-world analysts. The study acquires the user's

interpretation of display components, captures their cognitive processes as well as contextual knowledge, and quantitatively compares tabular versus graphical displays. An additional aspect of the study compares real-world analyst feedback with that of students, since students are the primary test subjects for academia developing visual displays for network monitoring. In this quantitative study, participants act as analysts and their job is to identify as many of the network threats on the simulated network provided. We will observe the participants' responses to the pattern-matching activity created within the game scenario. The design variables will be the distinct graphical layouts. The response variables are true-positive and false-positive rates of event identification, the time required for event identification, and the qualitative questionnaire. Results will help us understand which of the visual layouts is most effective for predicting cyber attacks. This will benefit network security analysts who defend the nation's networks.

2. Information Visualization

Automated systems requiring vigilant human insight are one potential solution to combat computer security threats. It is recommended that these systems incorporate a human in the diagnostic loop since his/her analytic skills far surpass that of computers (2). In general, support tools are needed to integrate intricate sense-making capabilities with the ability of these automated systems to process vast quantities of data (4). Information visualization is defined (28) as a computer-supported, interactive visual representation of data to amplify cognition. Information visualization is one such method that shows great potential for supporting computer security work in that it provides the human security analysts with better tools to discover patterns, detect anomalies, identify correlations, and communicate findings, all while keeping the human in the diagnostic loop. Information visualization can be used for exploration discovery, decision-making, and communication of complex ideas, and it helps to deal with processing the influx of data. This is an interactive method used to represent abstract data when compared to other data graphics. Information visualization tools allow the user to adjust the display in order to gain a more meaningful understanding of the data being presented. Mapping the data spatially in a meaningful matter is the most important and challenging part to making an effective information visualization (4). At the core of information visualization is the goal of amplifying cognition, the intellectual processes in which information is obtained, transformed, stored, retrieved, and used (3). Robust information visualization tools that implement the importance of keeping humans in the loop take advantage of the power of the human perceptual and cognitive processes in solving computer security problems.

2.1 Visual Representation in Table Form

Analysts are used to, and most times prefer, tabular displays. Tabular displays originate from spreadsheet techniques that provide a structured, intuitive, and powerful interface for investigating information visualizations of multidimensional datasets (5). Mathematicians and

statisticians have long used tables of sine, cosine, and confidence probabilities. Previously, the invention of the VisiCalc numerical spreadsheet in 1979 fueled the adoption of usage with personal computers (6). Statisticians have examined visualizing higher dimensional point sets by a table of projections. For example, one multivariate analysis tool is the scatter matrix, which is a table of scatter plots (7). Since the early 1980s, visualization researchers have applied similar ideas, but in different ways, to produce a table of views of a single dataset (8, 9). These approaches represent a largely static tabular approach to the data, but some interactivity is present, such as rotations, translation, and zooming. There are several distortion presentation techniques based on a tabular layout (10) such as Document Lens (11), fish-eye views (12, 13), stretching rubber sheets (14). Overall, the advantages for analysts using the tabular layout are that it is familiar, flexible, easily configurable, and excellent for interactive comparison tasks (5).

2.2 Visual Representation in Graphical Form

We recommend using graphical representations to illustrate network activity and relationships among network components. This study is an approach toward providing analysts with enhanced visual displays that will ease their difficult task of integrating technical knowledge with contextual knowledge under severe constraints. Much research has been done to identify and develop external aids that enhance cognitive abilities for end-users. Visualization, itself, has been identified as a necessary and effective technology for network security, particularly, with intrusion detection systems (IDS) (26). While information visualization remains a novelty for some users, who struggle to use the graphics effectively, this study's suggested graphical representations of network data highlight components, patterns, relationships, and features that increase the utility of user displays and the likelihood of adoption by industry.

Various workflow visualization tools are available to help users track their analysis, reuse effective workflows, and test hypotheses (1). However, the need still exists for analysts to improve communication and performance, explore deeper into certain network attacks, and investigate suspicious activities within a network (27). Some past visualization techniques have contributed to better visual displays for end-users:

Flow-Based approaches (HistoryFlow, ThemeRiver, TimeWheel, Wormplots) (28–31)

- Glyph-Based approaches (32, 33)
- Circle Segment (34, 35)

Our approach to addressing analyst's needs for visualization information differs from previous works in two ways: the tools used and the focus of what is being visualized. In our case, we plan to turn visual overload into an opportunity by enhancing the visual displays used by analysts into a more effective tool. Traditionally, analysts are used to working with tabular displays for conducting their tasks. Other types of displays, such as a graphical display, may provide more insight into big data inter-relationships, patterns, and finding areas of interest. In response, we have designed an experiment using cyber-network data to test the effectiveness for

communicating suspicious activity on a computer network through visual displays. In the study, participants act as analysts, and their job is to identify as many as possible of the intrusion attacks and intrusion attack attempts on the tabular and graphical displays provided. The design variables will be several distinct graphical layouts. The response variables are true-positive and false-positive rates of event identification, the time required for event identification, and a qualitative questionnaire. Results will help us understand which of the visual layouts is most effective for predicting cyber attacks. This will benefit network security analysts who defend the nation's networks.

3. User Studies

Evaluating scientific visualization techniques is a longstanding challenge (15–17). Similarly, the field of information visualization has a strong tradition in pioneering research in evaluation techniques (18–20). User studies often rely on timing and accuracy information collected during the study, coupled with subjective user surveys given after the experiment is completed. This combination of empirical measurement with a subjective questionnaire is designed to assess the efficacy of a visualization technique with respect to related methods. However, the analysis of user evaluation studies remains difficult. These challenges are often compounded by the limited empirical data acquired during the study. Beyond the specific details of the many user study experiments, they all share a common goal: to assess the strengths and weaknesses inherent to a visualization technique or system. Incorporating as many objective measures as possible into the experiment not only provides a more robust analysis, but also mitigates subjectivity often introduced by users' preferences, biases, and retrospection. In this position paper, we review traditional evaluation techniques that consist of data gleaned from system logging. We then outline evaluation methods using physiological measures for the assessment of scientific visualization efficacy.

Due to the nature of today's complex scientific data, simply displaying all available information does not adequately meet the demands of domain scientists. Determining the best use of visualization techniques is one of the goals of scientific visualization evaluations. The types of improvements offered by the method being studied dictate evaluation methods. Some evaluations are concerned primarily with technological improvements, such as rendering speed or the management of large data. User studies have been used to evaluate everything from aircraft cockpits (21) and surgical environments (22) to visualization methods (23). Evaluating visualization methods that focus on human factors often employ user studies or expert evaluations to determine their effects on interpretation and usability. An expert assessment takes advantage of knowledgeable users to enable more poignant analysis of use cases, and these experts also bring their own preconceptions and preferences that can skew studies. Traditional evaluation methods provide mechanisms to gauge aspects of visualizations or environment.

Unfortunately, experiments using surveys to measure user experience introduce subjectivity and bias from the users. Subjectivity in user responses may be partially mitigated using questionnaires developed with the Likert Scale (24). Subjectivity in evaluation may provide important insights into how users interact with the systems being studied. However, subjective measures do not help answer questions regarding how effective a method is at eliciting insight from a dataset. This is a primary purpose of visualization. Our goal and purpose is to use this project as an empirical study to examine the cognitive aspects of visual displays, with the goal of identifying components and representations that most effectively aid the computer network analyst in interpreting the underlying activity in a network sample. Results from the study are helpful to understand the potential and limitations of the suggested visual displays attempting to aid analysts' needs to better achieve their tasks.

3.1 Study Development

Step 1: Performed literature review on the following topic areas:

- Existing Visual Representations:
 - What current visual representations exist that could be applied to analysts' displays?
- Visualization Tools:
 - What visualization tools are currently being used for analysts' displays?
 - What is it about the tools that work for analysts and what analysts needs remain unmet by these tools?
 - Are there other visualization tools not specific to the network domain that could be of use for analysts display visualization needs?
- Existing User Studies
 - What studies have been done with visualization displays?
 - What studies have tested analysts' displays?
- Consider New Methods for Displays
 - What visualizations have been effective on displays used in other domains (medical field, biology field, etc.)?

Step 2: Completed the following Human Factors Trainings:

- The Principal Investigators (PIs) had to complete the Collaborative Institutional Training Initiative (CITI) at <https://www.citiprogram.org> and score at least 80% on each exam.
- The Participants Investigators (PIs) had to also complete the National Institutes of Health (NIH) at <http://phrp.nihtraining.com/> and pass each exam.

Step 3: Developed a protocol for the study by the following:

- The principal investigators held meetings to discuss parameters and theory of the study:
Research resulted in an experimental design for the study. Generally, the study is broken into two parts: a preliminary study that uses graphical methods to present network

information to users on a display, and a follow-up study that uses a game scenario to present the same displays but allows user interaction with the visual representations in the game for data exploration of interesting features. The preliminary study compares graphical methods to tabular displays typically used in real network analyst environments. There are two phases for the study. The objective of Phase 1 is to evaluate how the users' abilities to detect network "intrusions and possible intrusions" is affected by the three display strategies. In the preliminary study in Phase 1, information will be presented in static displays, and in the follow-up study in Phase 1, information will be presented by the CyFall game. In Phase 2, an emphasis will be placed on understanding analysts' cognitive processes. Phase 2 uses a reasoning support system developed by Penn State University to assist the analyst in formulating hypotheses about the state of a network. The steps the analysts engage in to formulate and discard hypotheses will be recorded.

- Collaboration:

ARL's Computational and Information Sciences Directorate (CISD) teamed up with the Human Research and Engineering Directorate (HRED) to form and conduct the study. We invited Morgan State University (MSU) student subjects as participants in the study; contracted Stony Brook University to manufacture software for the cyber-network game scenario, CyFall; and involved Pennsylvania State University (PSU), who produced the trace software to capture analysts' cognitive processes.

- Subjects:

This study compares the results of expert analysts with that of university students since university students are the primary test subjects for academia developing visual displays for network monitoring. The expert analysts come from ARL's Sustaining Base Network Assurance Branch (SBNAB) team at both the Adelphi Laboratory Center (ALC) and Aberdeen Proving ground (APG) locations. The university students come from MSU.

- Formulating Questionnaires:

- Demographic questions were created to establish the census of the subjects participating in the study. Questions inform of confidence level with pattern-matching activities, prior experience in analysts tasks, and into what populations they fall (male/female, age, etc.).
- Pre-task questions were created to measure the subject's subjective perception of representations.
- Post-task questions were created to determine performance satisfaction, and to gather the overall aptitude of the tools used, visual representations seen, game environment response, and special user insights.

We implemented the questionnaires in a Web-based open source survey application called LimeSurvey (36) for both the preliminary and follow-up studies. See figures 1 and 2 for

screenshots of question in LimeSurvey. The entire questionnaire sets may be found in the full protocol located in appendix A.

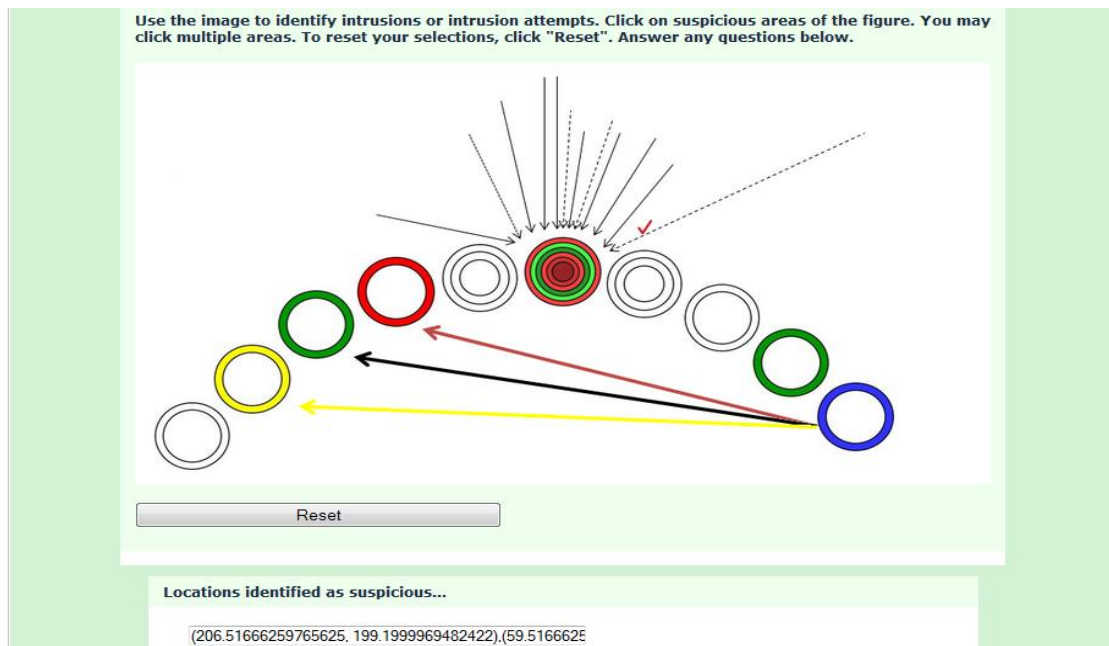


Figure 1. A screenshot of the node-link graphical representation of computer network alerts. The user is asked here to determine regions of the visualization that imply intrusions (True Positive [TP]) and intrusion attempts (False Positives [FP]) by clicking near a particular link or node.

Use the table to identify intrusions or possible intrusion attempts by checking the box in the associated row of data. You may tap a column heading to sort data by the values in that column or select a value from the drop-down box in each column to filter data. Answer any questions below.

Suspicious	Date	Time	ToolName	Protocol	SourceEn	SourceIP	SourcePo	DestEntit	DestIP	DestPort	Country	SrcCC	DstCC	AlertMes	AlertTraff
<input type="checkbox"/>	10/17/201	15:00:30	Snort	tcp	US1.2	10.234.111 52233	MD0.6	10.66.0.6	52030	MD	US	MD	MD	ET TROJAN [09 00 00]	
<input type="checkbox"/>	10/17/201	15:00:45	Snort	tcp	MD0.6	10.66.0.6 80	US1.2	10.234.111 52200	MD	MD	US	MD	MD	ET TROJAN PONG [3a]	
<input checked="" type="checkbox"/>	10/17/201	15:00:50	Snort	tcp	US1.2	10.234.111 81	MD0.6	10.66.0.6	52330	MD	US	MD	MD	ET TROJAN [09 00 00]	
<input type="checkbox"/>	10/17/201	15:01:10	Snort	tcp	US1.2	10.234.111 52001	MD0.6	10.66.0.6	80	MD	US	MD	MD	ET TROJAN GET [20/]	
<input type="checkbox"/>	10/17/201	15:01:35	Snort	tcp	US1.2	10.234.111 52001	MD0.6	10.66.0.6	80	MD	US	MD	MD	ET TROJAN	
<input checked="" type="checkbox"/>	10/17/201	15:59:05	Aggregate	tcp	US1.3	10.234.111 *	US1.0	10.250.10. 80	US	US	US	US	US	High traffic *	
<input type="checkbox"/>	10/17/201	15:20:12	Snort	tcp	US1.5	10.234.111 51119	US1.0	10.250.10. 80	US	US	US	US	US	ET TROJAN GET /instal	
<input type="checkbox"/>	10/17/201	15:59:00	Aggregate	tcp	US1.5	10.234.111 *	US1.0	10.250.10. 80	US	US	US	US	US	High traffic *	
<input type="checkbox"/>	10/17/201	15:32:15	Snort	tcp	US.55	10.234.111 80	US1.0	10.250.10. 80	US	US	US	US	US	ET TROJAN GET /instal	
<input checked="" type="checkbox"/>	10/17/201	15:45:00	Aggregate	tcp	US.55	10.234.111 *	US1.0	10.250.10. 80	US	US	US	US	US	High traffic *	
<input type="checkbox"/>	10/17/201	15:45:50	Snort	tcp	US0.2	10.123.100 55500	US1.0	10.250.10. 80	US	US	US	US	US	ET TROJAN GET /instal	
<input type="checkbox"/>	10/17/201	15:59:00	Aggregate	tcp	US0.2	10.123.100 *	US1.0	10.250.10. 80	US	US	US	US	US	High traffic *	

Checked rows are...

3,6,10

Figure 2. A screenshot of the tabular representation of computer network alerts. The user is asked here to determine which alert messages in the table imply intrusions (True Positive [TP]) and intrusion attempts (False Positives [FP]) by clicking the checkboxes in the 'Suspicious' column.

Step 4: Review Process

- The completed protocol was then sent for technical review by team lead, supervisor, external reviewers, branch chief, division chief, and the human factors administrator.

During the review process, there were several revisions made to the protocol. Changes included breaking the study up into two parts: a preliminary study to obtain initial visual display feedback from users, and a follow-up study to incorporate the visual displays into a cyber-network game scenario similar to real network analysts tasks. The approved protocol for the study with project number ARL 13-050 is attached as appendix A. See figure 3 for an overview of the development for this study.

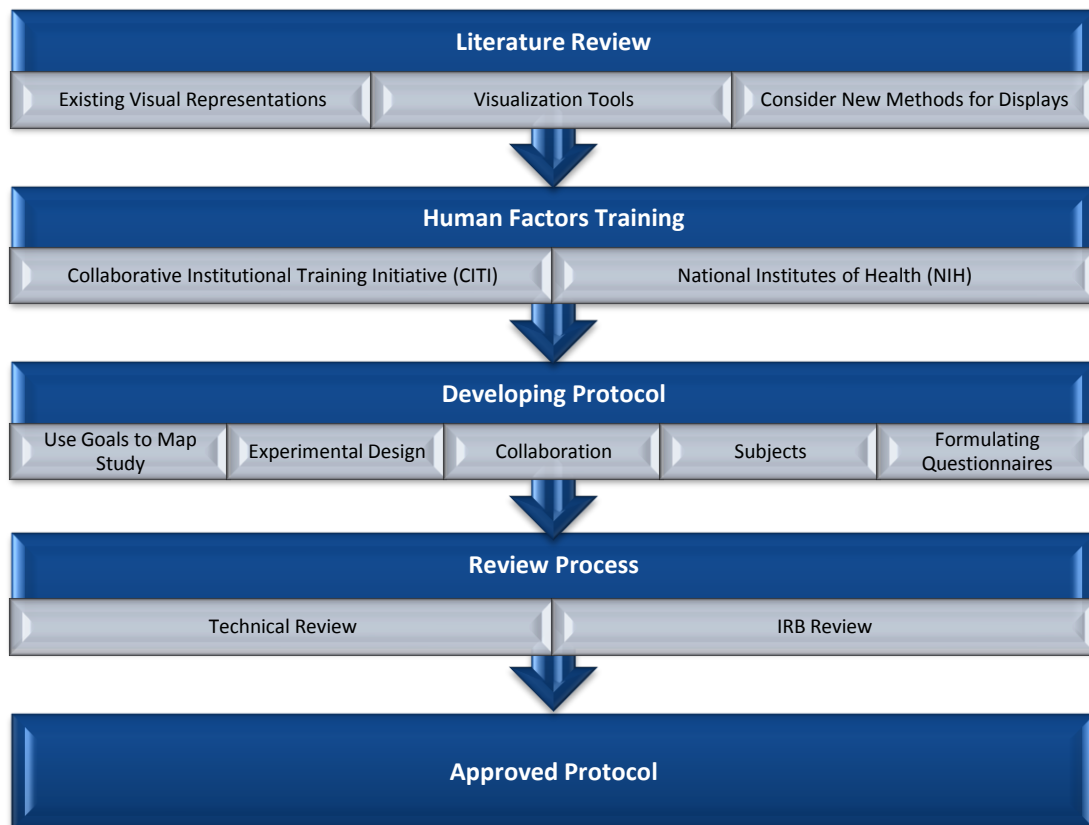


Figure 3. Organization chart of an overview of the development for this study.

3.2 Experimental Design

Two goals for this study frame the experimental design. The first design we call Phase I includes three visual displays: Tabular, Parallel Coordinates, and Node-Link where we have examined their cognitive aspects to further identify components and representations that most effectively aid the CND analyst in interpreting the underlying activity in a network data sample. The second design we call Phase II uses a tool to capture the analyst's cognitive reasoning process. Phase I investigates the various representations, and Phase II makes use of a tool designed to understand the process by which analysts perform their analysis. A laptop will display several figures

Alerts.csv - OpenOffice.org Calc													
File Edit View Insert Format Tools Data Window Help													
Libration Sans 10													
A	B	C	D	E	F	G	H	I	J	K	L	M	N
	INCIDENT	Date	Time	Location	Protocol	Source	SourcePort	ClientIP	ClientPort	Country	AlertMessage/Method	Alert/Task/Script	Notes
1	1	10/17/12	03:00:30 PM	Smart	tcp	10.234.111.2		52300.60	52300.60	ET TROJAN Backdoor Connect to Controller"	705 00 00 [aj] [74]"		200872.4
2	1	10/17/12	03:00:35 PM	Smart	tcp	10.66.66		52300.60	52300.60	ET TROJAN Backdoor Response from Controller"	705 00 00 [BQ]"		200872.4
3	1	10/17/12	03:00:40 PM	Smart	tcp	10.234.111.2		42200.66	42200.66	ET TROJAN Backdoor Connect to Controller (PING PONG)"	705 00 00 [aj] [74]"		200872.4
4	1	10/17/12	03:00:45 PM	Smart	tcp	10.66.66		42200.66	42200.66	ET TROJAN Backdoor Response from Controller (PING PONG)"	705 00 00 [BQ]"		2009129.2
5	1	10/17/12	03:00:50 PM	Smart	tcp	10.234.111.2		42200.66	42200.66	ET TROJAN Backdoor Response from Controller"	705 00 00 [aj] [74]"		2009129.2
6	1	10/17/12	03:01:00 PM	Smart	tcp	10.234.111.2		42200.66	42200.66	ET TROJAN Backdoor Response from Controller"	705 00 00 [aj] [74]"		2009129.2
7	1	10/17/12	03:01:05 PM	Smart	tcp	10.234.111.2		42200.66	42200.66	ET TROJAN Backdoor Response from Controller"	705 00 00 [aj] [74]"		2009129.2
8	1	10/17/12	03:01:10 PM	Smart	tcp	10.234.111.2		42200.66	42200.66	ET TROJAN Backdoor Response from Controller"	705 00 00 [aj] [74]"		2009129.2
9	1	10/17/12	03:01:15 PM	Smart	tcp	10.234.111.2		42200.66	42200.66	ET TROJAN Backdoor Response from Controller"	705 00 00 [aj] [74]"		2009129.2
10	1	10/17/12	03:01:20 PM	Aggregate	tcp	10.234.111.2		42200.66	42200.66	ET TROJAN Backdoor Response from Controller"	705 00 00 [aj] [74]"		2009129.2
11	1	10/17/12	03:01:25 PM	Aggregate	tcp	10.234.111.2		42200.66	42200.66	ET TROJAN Backdoor Response from Controller"	705 00 00 [aj] [74]"		2009129.2
12	1	10/17/12	03:01:30 PM	Aggregate	tcp	10.234.111.2		42200.66	42200.66	ET TROJAN Backdoor Response from Controller"	705 00 00 [aj] [74]"		2009129.2
13	1	10/17/12	03:01:35 PM	Aggregate	tcp	10.234.111.2		42200.66	42200.66	ET TROJAN Backdoor Response from Controller"	705 00 00 [aj] [74]"		2009129.2
14	1	10/17/12	03:01:40 PM	Aggregate	tcp	10.234.111.2		42200.66	42200.66	ET TROJAN Backdoor Response from Controller"	705 00 00 [aj] [74]"		2009129.2
15	1	10/17/12	03:01:45 PM	Aggregate	tcp	10.234.111.2		42200.66	42200.66	ET TROJAN Backdoor Response from Controller"	705 00 00 [aj] [74]"		2009129.2
16	1	10/17/12	03:01:50 PM	Aggregate	tcp	10.234.111.2		42200.66	42200.66	ET TROJAN Backdoor Response from Controller"	705 00 00 [aj] [74]"		2009129.2
17	1	10/17/12	03:01:55 PM	Aggregate	tcp	10.234.111.2		42200.66	42200.66	ET TROJAN Backdoor Response from Controller"	705 00 00 [aj] [74]"		2009129.2
18	1	10/17/12	03:02:00 PM	Aggregate	tcp	10.234.111.2		42200.66	42200.66	ET TROJAN Backdoor Response from Controller"	705 00 00 [aj] [74]"		2009129.2
19	1	10/17/12	03:02:05 PM	Aggregate	tcp	10.234.111.2		42200.66	42200.66	ET TROJAN Backdoor Response from Controller"	705 00 00 [aj] [74]"		2009129.2
20	1	10/17/12	03:02:10 PM	Aggregate	tcp	10.234.111.2		42200.66	42200.66	ET TROJAN Backdoor Response from Controller"	705 00 00 [aj] [74]"		2009129.2
21	1	10/17/12	03:02:15 PM	Aggregate	tcp	10.234.111.2		42200.66	42200.66	ET TROJAN Backdoor Response from Controller"	705 00 00 [aj] [74]"		2009129.2
22	1	10/17/12	03:02:20 PM	Aggregate	tcp	10.234.111.2		42200.66	42200.66	ET TROJAN Backdoor Response from Controller"	705 00 00 [aj] [74]"		2009129.2
23	1	10/17/12	03:02:25 PM	Aggregate	tcp	10.234.111.2		42200.66	42200.66	ET TROJAN Backdoor Response from Controller"	705 00 00 [aj] [74]"		2009129.2
24	1	10/17/12	03:02:30 PM	Aggregate	tcp	10.234.111.2		42200.66	42200.66	ET TROJAN Backdoor Response from Controller"	705 00 00 [aj] [74]"		2009129.2
25	1	10/17/12	03:02:35 PM	Aggregate	tcp</								

9

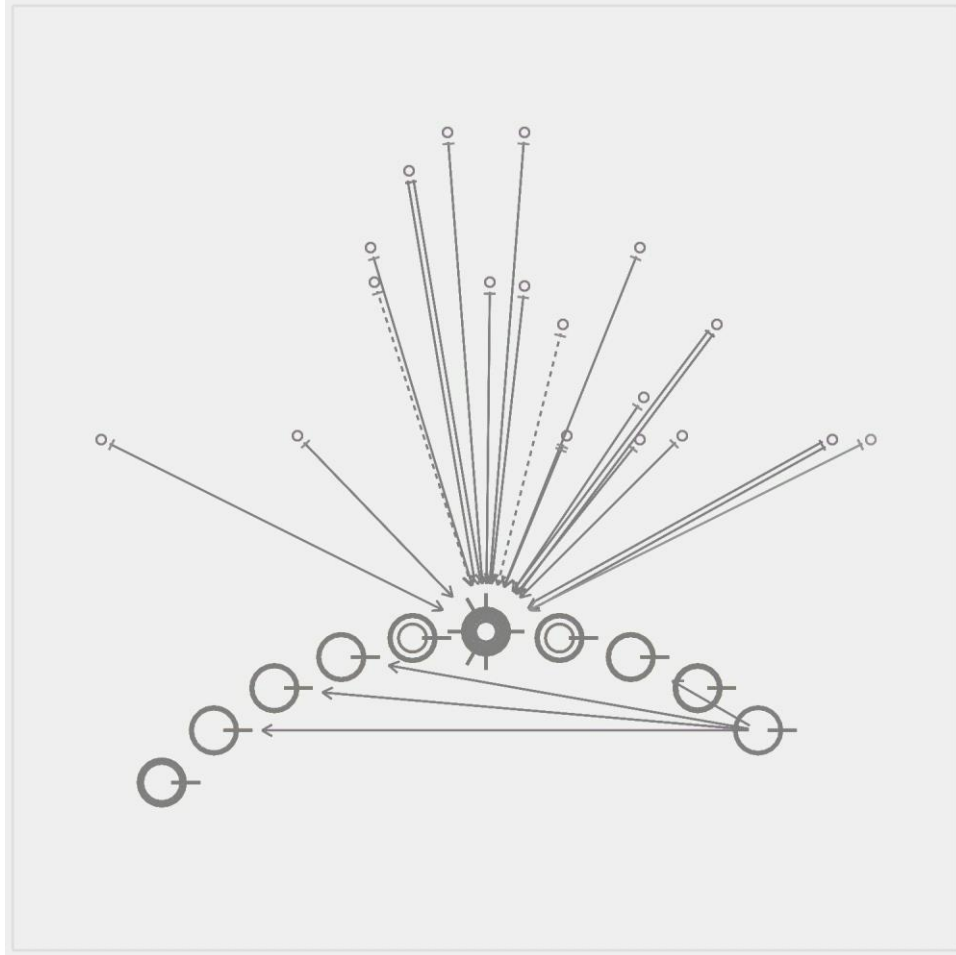


Figure 6. Node-link Display, representation C.

Hardware is used to conduct the experiment simultaneously with two participants. During the preliminary and follow-up studies, we will apply full randomization of the test subjects using the following possible sequences of the subjects viewing the visual displays on the hardware:

- Group 1: $A \rightarrow B \rightarrow C$
- Group 2: $A \rightarrow C \rightarrow B$
- Group 3: $B \rightarrow A \rightarrow C$
- Group 4: $B \rightarrow C \rightarrow A$
- Group 5: $C \rightarrow A \rightarrow B$
- Group 6: $C \rightarrow B \rightarrow A$

Thus, two subjects will perform each ordering of visual displays. While not statistically significant, this should begin to identify any impact of the ordering on performance, which, itself, may aid in training of future analysts. The fabricated dataset used for the visual displays and game was generated using threats from ThreatExpert (37). The threats were derived from the “Index of Open Snort 2.9.0 Rules” (38), which is publically available. See figure 7 for an overview of the experiment design.

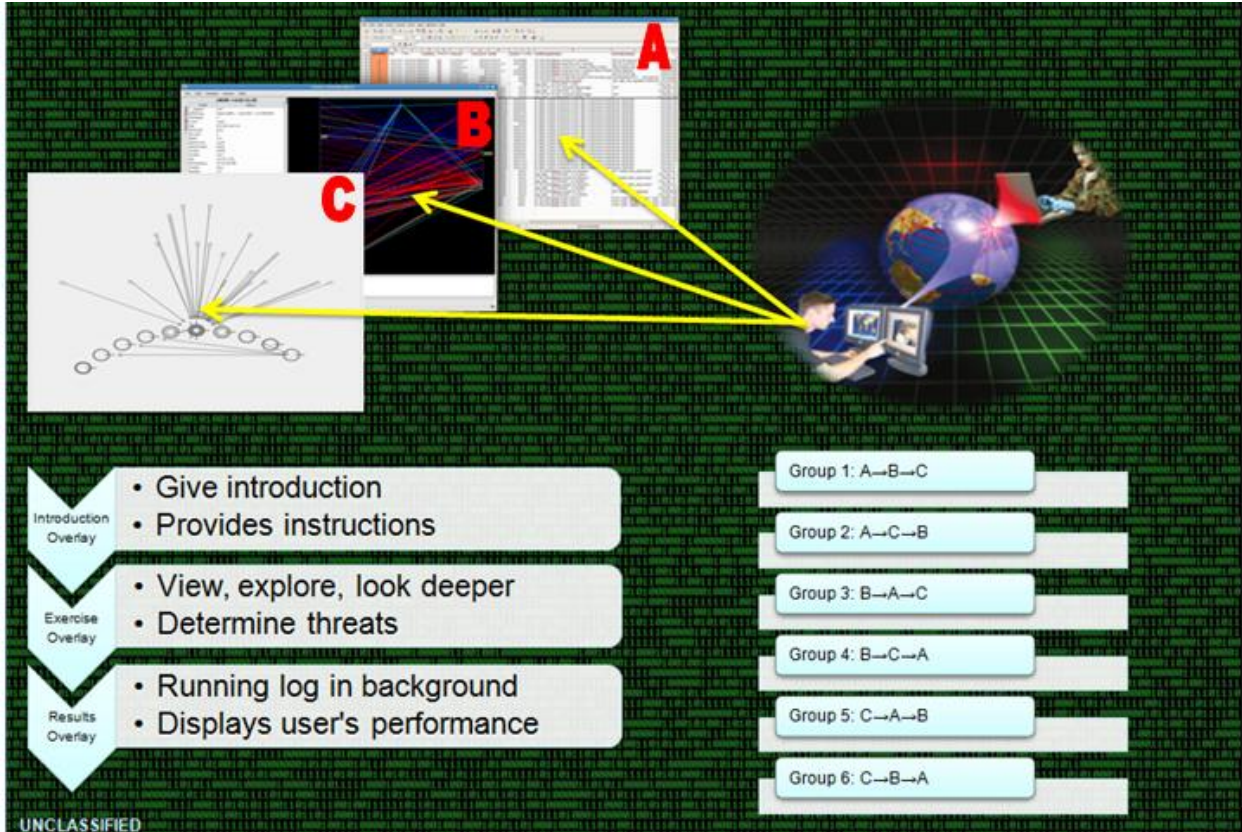


Figure 7. Experiment Design Overview.

4. Future Work

With a complete and approved protocol, we can now begin the preliminary and follow-up studies for the next fiscal year. The goal is to conduct the study with both the expert analysts and student users, collect the data, and analyze (accuracy, error rate, time, and quantitative questionnaires) the results. We plan to submit a technical report of our findings and to publish a paper for a conference or journal.

5. References

1. Singh, Ankit, et al. Supporting the Cyber Analytic Process Using Visual History on Large Displays. *Proceedings of the 8th International Symposium on Visualization for Cyber Security*, ACM, 2011.
2. Allen, J.; Christie, A.; Fithen, W.; McHugh, J.; Pickel, J.; Stoner, E. *State of the Practice of Intrusion Detection Technologies*; Tech. Rep. CMU/SEI-99-TR-028; Carnegie Mellon University/Software Engineering Institute, 1999.
3. Card, S. K. *Information Visualization*; In: Jacko, J.A., Sears, A. (eds.) *The Human Computer Interaction Handbook*, pp.544–582. Lawrence Erlbaum Associates, Mahwah, NJ, 2003.
4. Goodall, John R. Introduction to Visualization for Computer Security. *VizSEC 2007*, Springer Berlin Heidelberg, 1–17, 2008.
5. Chi, EH-H., et al. A Spreadsheet Approach to Information Visualization. *Information Visualization, 1997. Proceedings., IEEE Symposium on*. IEEE, 1997.
6. Brown, P. S.; Gould, J. D. An Experimental Study of People Creating Spreadsheets. *ACM Transactions on Office Information Systems* **July 1987**, 5 (3), 258–272.
7. Dayhoff, M. O.; Schwartz, R. M.; Orcutt, B. C. A Model of Evolutionary Change in Proteins. In M. O. Dayhoff, editor, *Atlas of Protein Sequence and Structure*, Vol. 5, Suppl. 3, chapter 22, pages 345–352, National Biomedical Research Foundation, 1978.
8. Anupam, V.; Dar, S.; Leibfried, T.; Petajan, E. Dataspace: 3-D Visualizations of Large Databases. In *IEEE Information Visualization Symposium*, pages 82–88, 144, 145, 1995.
9. Van Wijke, J.; Van Liere, R. Hyperslice: Visualization of Scalar Functions of Many Variable. In *IEEE Visualization '91*, pages 119–125, Los Altos, CA, 1991, IEEE CS Press.
10. Leung, Y. K.; Apperley, M. D. A Review and Taxonomy of Distortion-Oriented Presentation Techniques. *ACM Transactions on Computer-Human Interaction* **1994**, 1 (2), 126–160.
11. Robertson, G. G.; Mackinlay, J. D. The Document Lens. In *Proceedings of the ACM SIGGRAPH Symposium on User Interface Software and Technology, Visualizing Information*, pages 101–108, 1993.
12. Furnas, G. W. Generalized Fisheye Views. In *Proceedings of ACM CHI'86 Conference on Human Factors in Computing Systems, Visualizing Complex Information Spaces*, pages 16–23, 1986.

13. Sarkar, M.; Brown, M. H. Graphical Fisheye Views of Graphs. In *Proceedings of ACM CHI'92 Conference on Human Factors in Computing Systems, Visualizing Objects, Graphs, and Video*, pages 83–91, 1992.
14. Sarkar, M.; Snibbe, S. S.; Tversky, O. J.; Reiss, S. P. Stretching the Rubber Sheet: A Metaphor for Visualizing Large Layouts on Small Screen. In *Proceedings of the ACM SIGGRAPH Symposium on User Interface Software and Technology, Visualizing Information*, pages 81–91, 1993.
15. Acevedo, D.; Jacson, C.; Drury, F.; Laidlaw, D. Using Visual Design Experts in Critique-Based Evaluation of 2d Vector Visualization Methods. *IEEE Transactions on Visualization and Computer Graphics* **2008**, 14 (4), 877–884.
16. Acevedo, D.; Laidlaw, D. Subjective Quantification of Perceptual Interactions Among Some 2d Scientific Visualization Methods. *IEEE Transactions on Visualization and Computer Graphics* **2006**, 12 (5), 1133–1140.
17. Kosara, R.; Healey, C. G.; Interrante, V.; Laidlaw, D. H.; Ware, C. Thoughts on User Studies: Why, How and When. *IEEE Computer Graphics and Applications* **2003**, 23 (4), 20–25.
18. Shneiderman, B.; Plaisant, C. Strategies for Evaluating Information Visualization Tools: MILCS. *Proc. Of AVI Workshop BELIV 2006*, pages 1–7, 2006.
19. Riche, N. Beyond System Logging: Human Logging for Evaluating Information Visualization. *Proc. Of SIGCHI Workshop SELIV 2010*, 2010.
20. Carpendale, S. *Evaluating Information Visualizations*; In A. Kerren, J. Stasko, J. D. Fekete, and C. North, editors, *Information Visualization*, volume 4950 of *Lecture Notes in Computer Science*, pages 19–45. Springer Berlin / Heidelberg, 2008.
21. Sarter, N. B.; Woods, D. D. Pilot Interaction With Cockpit Automation II: An Experimental Study of Pilots' Model and Awareness of the Flight Management System. *Int'l J. of Aviation Psychology* **1994**, 4 (1), 1–28.
22. Reitinger, B.; Bornik, A.; Beichel, R.; Schmalstieg, D. Liver Surgery Planning Using Virtual Reality. *IEEE Computer Graphics and Applications* **2006**, 26, 36–47.
23. Laidlaw, D. H.; Kirby, R. M.; Jackson, C. D.; Davidson, J. S.; Miller, T. S.; Da Silva, M.; Warren, W. H.; Tarr, M. J. Comparing 2d Vector Field Visualization Methods: A User Study. *IEEE Transactions on Visualization and Computer Graphics* **2005**, 11 (2), 59–70.
24. Likert, R. A Technique for the Measurement of Attitudes. *Archives of Psychology* **1932**, 140, 1–55.

25. Becker, Richard A., et al. Dynamic Graphics for Network Visualization. *Visualization, 1990. Visualization'90. Proceedings of the First IEEE Conference on. IEEE*, 1990.
26. Cox, Kenneth C.; Eick, Stephen. G. Case Study: 3D Displays of Internet Traffic. In *Information Visualization Symposium*, Atlanta, Georgia, October 1995.
27. Nyarko, Kofi, et al. Network Intrusion Visualization with NIVA, an Intrusion Detection Visual Analyzer with Haptic Integration. *Haptic Interfaces for Virtual Environment and Teleoperator Systems, 2002. HAPTICS 2002. Proceedings. 10th Symposium on. IEEE*, 2002.
28. Harve, Susan; Hetzler, Beth; Nowell, Lucy. Themeriver: In Search of Trends, Patterns, and Relationships. In *IEEE Transactions on Visualization and Computer Graphics*, volume 8, pages 9–20, IEEE Computer Society Press, 2002.
29. Tominski, Christian; Abello, James; Schumann, Heidrun. Axes-Based Visualizations with Radial Layouts. In *Proceedings of the 2004 ACM Symposium on Applied Computing*, pages 1242–1247, ACM Press, 2004.
30. Treinish, Lloyd; Silver, Deborah. Worm Plots. *IEEE Computer Graphics and Applications* **1997**, 17, 17–20.
31. Viegas, Fernanda B.; Wattenberg, Martin; Dave, Kushal. Studying Cooperation and Conflict Between Authors with History Flow Visualizations. In *Proceedings of 2004 conference on Human Factors in Computing Systems*, volume 6, pages 575–583, ACM Press, 2004.
32. Chuah, Mei C.; Eick, Stephen G. Information Rich Glyphs for Software Management Data. *Computer Graphics and Applications, IEEE* **1998**, 18 (4), 24–29.
33. Cviz website [online]. June 2007. Available from: <http://www.alphaworks.ibm.com/formula/CViz> [cited 2007-06-23].
34. Kiem, Daniel A. Visual Techniques for Exploring Databases. In *Tutoring Notes: Third International Conference on Knowledge Discovery and Data Mining*, pages 1–121, AAAI Press, 1997.
35. Carlis, John V.; Konstan, Joseph A. Interactive Visualization of Serial Periodic Data. In *Proceedings of the 11th Annual ACM Symposium on User Interface Software and Technology*, pages 29–38, ACM Press, 1998.
36. Schmitz, C. Limesurvey-the Open Source Survey Application. Hamburg [cited 31.03. 2009] Available from: <http://www.Limesurvey.org>. External link (2009).
37. ThreatExpert - Automated Threat Analysis. ThreatExpert.com, 2009. Web. 13 Sept. 2013.
38. Emerging Threats. (n. d.). emerging-all.rules. Retrieved from Emerging Threats: <http://rules.emergingthreats.net/open/snort-2.9.0/emerging-all.rules>.

Appendix A. Approved Protocol Project#: ARL 13-050

HUMAN SUBJECT RESEARCH PROTOCOL U.S. Army Research Laboratory Adelphi Laboratory Center, MD 20783-1197

Title

Evaluation of the Presentation of Network Data via Visualization Tools for Network Analysts

Co-Principal Investigator

Name: Renée E. Etoty

Directorate: Computational and Information Sciences Directorate

Division: Network Science Division

Branch: Network Security Branch

Team: Advanced Intrusion Detection

Phone Number: 301-394-1835

Email: renee.e.etoty.civ@mail.mil

Co-Principal Investigator

Name: Robert F. Erbacher

Directorate: Computational and Information Sciences Directorate

Division: Network Science Division

Branch: Network Security Branch

Team: Advanced Intrusion Detection

Phone Number: 301-394-1674

Email: robert.f.erbacher.civ@mail.mil

Associate Investigator

Name: Christopher Garneau

Directorate: Human Research and Engineering Directorate

Division: Human Factors Integration Division

Branch: MANPRINT Methods and Analysis

Team: Tool Development

Phone Number: 410-278-5814

E-mail: christopher.j.garneau.civ@mail.mil

Location of Study

Building 204, Room 2F014

Adelphi Laboratory Center, MD

Building 459, Room 202 (System Assessment and Usability Laboratory)
Aberdeen Proving Ground, MD

Engineering Visualization Research Laboratory (EVRL)
Schaefer Engineering Building, Room 112
Morgan State University, Baltimore, MD

Abstract

The goal of security visualization is to help analysts increase the safety and soundness of our digital infrastructures by providing effective tools and workstations (16). Analysts have a difficult task characterized by the need to integrate technical knowledge with contextual knowledge under severe constraints. In our case, we plan to turn visual overload into an opportunity by enhancing the visual displays used by analysts into a more effective tool. Traditionally, analysts are used to working with tabular displays for conducting their tasks. Other types of displays such as a graphical display may provide more insight into big data inter-relationships, patterns, and finding areas of interest. In response, we have designed an experiment using cyber-network data to test the effectiveness for communicating suspicious activity on a computer network through visual displays. In the study, participants act as analysts and their job is to identify as many as possible of the intrusion attacks and intrusion attack attempts on the tabular and graphical displays provided. The design variables will be several distinct graphical layouts. The response variables are true positive and false positive rates of event identification, the time required for event identification, and a qualitative questionnaire. Results will help us understand which of the visual layouts is most effective for predicting cyber attacks. This will benefit network security analysts who defend the nation's networks.

Location of Research

We will conduct the research at Adelphi Laboratory Center (ALC), Aberdeen Proving Ground (APG), and Morgan State University (MSU).

Data Collection Dates

The data collection dates will take place 1 September 2013 through 1 October 2014.

Study Sponsor

U.S. Army Research Laboratory Computational and Information Sciences Directorate (ARL-CISD)

Research Background

Millions of data features can quantify the structure of complex cyber networks. However, information overload is a persistent problem existing in graphical layout techniques. In a graph, nodes represent objects under analysis and links represent the relationship between these elements. The drawing of nodes and their edges onto a two-dimensional surface is a difficult problem with no satisfying solution. Challenges arise in the representation of graphs visually due

to the complexity of the graphs and the difficulty in removing occlusion in 2D projections while representing the overarching ontology.

Tables work best for representing data when the presentation is used to look up or compare individual values, when precise values are required, and when the values involve multiple units of measure (17). Graphical representations, work best when the data presentation is used either to communicate a message that is contained in the shape of the data or to reveal the relationship among many values. Hence, graphs and tables are the two primary means to structure and communicate quantitative information. Surprisingly, not much work has been done in the cyber security domain to validate or disprove the effectiveness of either display type being used in the analysis of cyber security tasks. Goodall (19) is one of few whose work focused on comparing user performance by using two different tools designed to analyze captured network packet data. Traffic Network Visualization tool (TNV) was the visualization-based display tool used. It was compared to a textual-tabular-based display tool. In (19), TNV proved increased accuracy for well-defined tasks. They also mention a clear preference from their expert participants for the visual interface. While other graphical visualization tools have been developed and prove useful such as Koike and Ohno's SnortView (20) that used simple geometric shapes to indicate protocol and severity in two-dimensional grid relating source IP address to destination IP address and time, Goodall remains the only known approach comparing tabular and graphical displays validated by a user study. Our approach to enhancing analyst's needs for visual displays differs from Goodall in two ways, the tools used and the focus of what is being visualized. We initially use the MATLAB tool to house a tabular display that will be compared to several graphical displays and we focus on visualizing network traffic monitored by analysts rather than the correlated IDS output in (19). Our results will validate improved accuracy and the desire for visual displays that are more effective. Our study helps in creating a sound voice and reference for the cyber security domain concerning this matter. We agree with the authors (21) about the importance of grounding cyber security visualizations through user studies.

Ongoing research continues to look for new ways to provide decision-making opportunities that improve the effectiveness of cyber-security network analysts' activities. Effective visualization techniques can identify predictive features and reduce the dimensionality of both data and model while identifying relevant patterns (6). We aim to understand the underlying characteristics of effective cyber security monitoring such that we can minimize the information displayed to analysts. The ultimate goal for an effective display is to improve task performance enhancing situational awareness accuracy or decrease cognitive load. Contributing factors for cognitive load include perception, problem solving, and multi-tasking. We want to identify the salient features that analysts respond to best for each graphical layout of the visualization tool's environment. The relevance of the study is to help us better focus on aspects within the visual representation that may require cognition. Our informal hypothesis is that there will be better knowledge of analyst response to visual stimuli that will allow the generation of visual representations to maximize saliency of features of interest for network analysts.

Previous work on visualization for cyber security has focused on data analysis, event analysis, event identification, and situational awareness (7). These studies have proposed visual alternatives but none have been tested on significantly large network data sample sizes, equitable data simulations, or using expert network analysts. An example of such visual alternatives is from Kosara, et al.'s semantic depth of field, in which renderings strive to induce perceptual changes in the user (4). Tory and Möller (5) offer a thorough discussion of human factors in user study methods, and visualization design.

Evaluating visualization techniques can be a difficult task. The primary approaches include subjective feedback from domain experts and quantitative user studies. We will employ quantitative user studies as our primary approach in this study. There are several methods for measuring user response during visual user studies. This includes direct user manipulation, Electroencephalography (EEG) and Functional Magnetic Resonance Imaging (fMRI). We will employ direct user manipulation in which users directly respond to display elements while their time and performance are measured. EEG, a process of passively recording brain activity, is an alternative method for quantitatively evaluating visualization techniques. The measurements collected by EEG determine the amount of burden placed on an individual's cognitive resources. Anderson et al. (8) used this method for analysis of their visualization technique. Another method is fMRI, which is the process of detecting changes in blood oxygenation and flow that occur in response to neural activity (10). An active brain area consumes more oxygen and in response to the demand blood flow increases to that area. This method produces activation maps that show which parts of the brain are involved in a particular brain activity (10). Unfortunately, there are some disadvantages to using fMRI. It is expensive, clear images are only captured if the person being scanned remains completely still, and researchers are still uncertain of how fMRI really works. The disadvantages of using EEG are that it provides a view of overall brain activity, which is not specific to different areas of the brain and it requires attachment of electrodes to the subjects. We leave measuring cognitive load with EEG or fMRI to future studies.

Research Objective

This empirical study will examine the cognitive aspects of visual displays with the goal of identifying components and representations that most effectively aid the computer network analyst in interpreting the underlying activity in a network data sample. An additional objective is to capture the analyst's cognitive reasoning process via analysts recording their sequence of thoughts while conducting network defense tasks. The understanding obtained from the results in this study will allow for the generation of visual representations that maximize saliency of features of interest for network analysts and aid in building the foundation for science and theory of network-based intrusion detection. A preliminary study will gather preliminary results via subjects' response to generated visual representations of cyber-network data presented in capable environments such as the MATLAB tool. We also plan to conduct a follow-up study using a cyber-network attack game scenario to support our study's objectives. The game will identify

characteristics from the visual displays that are more effective for cyber analysis by using the results of student subjects versus expert subjects in a cyber-network game scenario.

Instrumentation and Facilities

Equipment

- At least two Laptops
- Installation/use of the visualization software (i.e., MATLAB etc.) on all hardware
- Installation/use of PSU trace tool on all laptops
- Installation/use of game software on all laptops (for follow-up study only)
- One projection machine

Safety Releases for Equipment or Apparatus

No safety releases are required.

Facility

We plan to conduct the studies at locations convenient to both network analysts and students. In particular, we plan to use the System Assessment and Usability Laboratory (SAUL) on the second floor of building 459 at APG, MD for subjects at that location and use room 2F014 in building 204 for subjects at the ALC location. There are several vacant rooms on the 2F00 hallway available for this study. SAUL and these rooms are good fits for this study because their primary function is to execute studies for software user interfaces.

Standard Operating Procedures for Courses or Facilities

There is currently no SOP on file for these facilities.

Materials, Tests, Tasks, and Stimuli

There are three questionnaires prepared for this study. Participants will sign the consent form and then take the surveys. The first is a “Demographic Questionnaire” which asks the subjects to provide background information and their level of experience related to the domain of the study (appendix B). The “Subjective Questionnaire” allows the subject to rate their experience with the game scenario manipulation of the graphical layout (appendix C). Last, a “Survey Questionnaire” asks the subject to rate their overall experience, identify their preferences associated with the different visual displays, and assess their use of the visualization tool (appendix D).

Tasks and Stimuli

A laptop will display several figures depicting network traffic using different graphical layout modes such as the one depicted in (appendix A). The participants will be given instructions on

how to interpret features of the visual displays provided for the preliminary study. Their task is to examine the intrusions and intrusion attempts highlighted by each visual display and to provide feedback on the effectiveness of communication on each representation of cyber-defense network data. We will use the hardware to conduct simultaneously the experiment with two participants.

During the preliminary study, we will make use of MATLAB, a high-level language and interactive environment for numerical computation, visualization, and programming to implement the designated visual display paradigms. MATLAB provides tools that enable gaining insight into data. Documents of explored analysis can be created and shared as reports or published MATLAB code (22). MATLAB allows access to data from files, other applications, databases, and external devices that may be read in via popular file formats such as Microsoft Excel; text or binary files; image, sound, and video files; and scientific files such as netCDF and HDF (22). The tool's ability to perform exploratory data analysis to uncover trends, test assumptions, and build descriptive models is one of the main reasons we selected it for use in this study.

During the follow-up study, we plan to use a program developed for ARL called CyFall. We will tell the participants that they are playing a game where they will be acting as a real cyber analyst. The participants' goal within the game is to try to detect and identify all of the intrusions and intrusion attempts on the network, as presented by the visual display. The participants will use the same visual displays from the preliminary study to identify and label as many of the correlated pieces of evidence that exist for each incident.

Subjects

The participant population will consist of analysts from ARL CISD located at ALC and APG as well as students from Morgan State University. The analysts are the individuals who either currently or in the past used related visual displays for the performance of their daily activities. This group will consist of eight to twelve participants. Subjects can only participate in this study if they are eighteen or older. For this preliminary study and the follow-up study, sixteen to twenty-four participants are sufficient. Subjects will be recruited by direct solicitation via personal and e-mail communication. There is no supervisory pressure to participate in this human research study. The subject is free to leave the study at any time. The follow up study will also be conducted at Morgan State University to compare student subjects with the expert analysts (ARL subjects).

Sample Size Justification

The preliminary and follow-up studies are initial usability tests. We will use the results to guide the design of actual visual displays to be used by network analysts. Therefore, only a few users are required. Sixteen to twenty-four participants will be recruited for this study. This number of participants is more than sufficient to gather qualitative feedback.

Compensation

Participation in the study is voluntary and there is no compensation provided for the subject's time or input. However, their contributions and results of the study may improve the quality of visual displays for ARL analysts, allowing them to identify features of interest more efficiently and effectively.

Subject Recruitment

Subject recruitment will be done by direct solicitation via word of mouth and email. We will specifically target analysts from the Sustaining Base Network Assurance Branch (SBNAB) within ARL CISD. Initial contact will be made to the SBNAB branch chief and he will notify us of arrangements on how to proceed with recruiting analysts from SBNAB. See Appendix J for the initial ARL recruitment email.

Experimental Design

As noted in the "Research Objective" section, the study is divided into two experimental efforts; a preliminary study intended to gain basic insights into the different representations and a more comprehensive follow-up study that uses a cyber-defense network game scenario. Each of these studies is divided into two phases: Phase I investigates the various representations and Phase II makes use of a tool designed to understand better the process by which analysts perform their analysis. Each of these components is described in this section.

Phase I:

For Phase I, both the preliminary and follow-up studies use post-experiment surveys in conjunction with timing and task-related data to form a foundation for additional statistical analysis. The following types of visual representations are used:

- A tabular sort-able display, see figure A-1

Figure A-1. Tabular Display, representation A.

- A colored parallel coordinate representation of alerts and normal traffic with a data inspector pane, see figure A-2

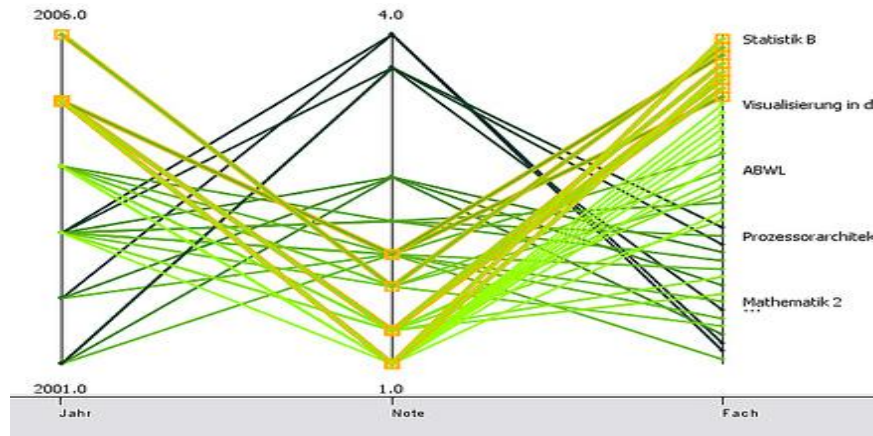


Figure A-2. Parallel Coordinates Display, representation B.

- A node-edge representation providing high-level situational awareness, see figure A-3

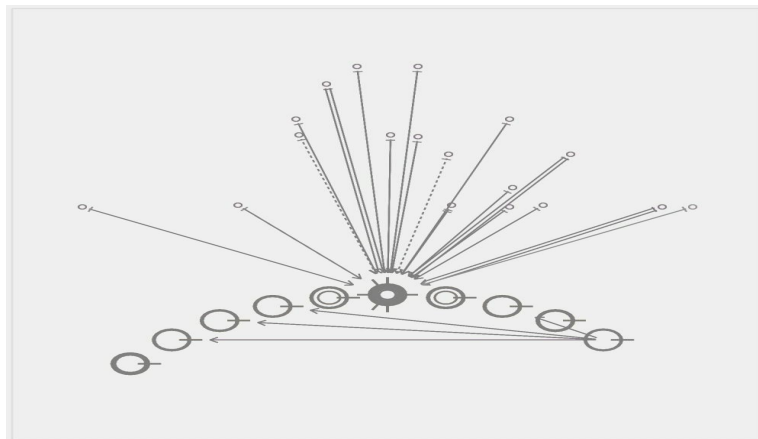


Figure A-3. Node-link Display, representation C.

Thus, there are at least three different ways to represent the same set of information. Design variables for both the preliminary and follow up studies are at least three graphical layouts. The response variables are true positive and false positive rates of event identification, the time required for event identification, and the qualitative questionnaires.

The Tabular Display provides data representing the exact attribution of the entire system typically presented in Microsoft Excel. An incident here is described as known bad senders, suspicious use of particular ports, and known patterns in the data packets. A list of alerts from a one-hour period is provided where the analyst can search for threats. Successful identification of a threat is considered a true positive (TP). Of course, it is a given that there will be some number of false alarm alerts we call false positives (FP). The analyst has a major task of differentiating the real threats from the false alarms. Figure A-4 for an example of what these alerts look like on the Tabular Display. For the follow-up study, the Tabular Display uses forensic techniques to collect and group evidence into what we call the victim system, which is the screen, on the display.

Date	Time	Protocol	SourceIP	SourcePort	DestIP	DestPort	Country	AlertMessageEmitted	AlertTrafficSnippet	RuleID	Class Type
10/17/12	15:00:01	tcp	10.123.100.2	21	10.234.111.3	21	US	"ET FTP USER login flowbit"	"USER "	2002850	not-suspicious
10/17/12	15:00:02	tcp	10.33.37.15	80	10.123.100.2	80	RO	ET WEB_SPECIFIC_APPS 1024 CMS standard.php page_include Parameter Remote File Inclusion	GET "/layouts/standard.php?" "page_include="	2009717	web-application-attack
10/17/12	15:00:03	tcp	10.234.111.4	21	10.99.0.2	21	GO	"ET FTP USER login flowbit"	"USER "	2002850	not-suspicious
10/17/12	15:00:20	tcp	10.123.100.2	21	10.234.111.3	21	US	"ET FTP HP-UX LIST command without login"	"LIST "	2002851	attempted-recon
10/17/12	15:00:20	tcp	10.200.25.80	80	10.123.100.9	80	US	ET WEB_SPECIFIC_APPS 35mm Slide Gallery imgdir Parameter Directory Traversal Attempt	GET "index.php?" "imgdir="	2010601	web-application-attack
10/17/12	15:00:23	tcp	10.77.124.60	80	10.123.100.7	80	SH	ET WEB_SPECIFIC_APPS AForum Remote Inclusion Attempt -- errormsg.php header	/common/errormsg.php? "header="	2003736	web-application-attack
10/17/12	15:00:23	tcp	10.22.138.125	80	10.234.111.206	21	ER	ET WEB_CLIENT PDF Name Representation Obfuscation of EmbeddedFile	obj "<<" "/" "EmbeddedFile" "#"	2011530	bad-unknown
10/17/12	15:00:23	tcp	10.22.146.41	80	10.234.111.80	80	ER	ET WEB_CLIENT PDF Name Representation Obfuscation of URL	obj "<<" "/" "URL" "#"	2011533	bad-unknown
10/17/12	15:00:24	tcp	10.22.169.141	80	10.234.111.103	80	ER	ET WEB_CLIENT PDF Name Representation Obfuscation of Pages	PDF- "/" "Pages" "#"	2011536	bad-unknown
10/17/12	15:00:29	tcp	10.22.151.84	80	10.123.100.9	80	ER	ET WEB_SPECIFIC_APPS 35mm Slide Gallery imgdir Parameter Directory Traversal Attempt	GET "index.php?" "imgdir="	2010601	web-application-attack
10/17/12	15:00:30	tcp	10.234.111.2	52233	10.66.0.6	52030	MD	"ET TROJAN Bifrose Connect to Controller"	"[09 00 00 9a] cc [74]"	2008273	trojan-activity
10/17/12	15:00:33	tcp	10.203.2.2	80	10.234.111.207	22	US	ET WEB_CLIENT PDF Name Representation Obfuscation of EmbeddedFile	obj "<<" "/" "EmbeddedFile" "#"	2011530	bad-unknown

Figure A-4. Alerts Example in a Tabular Display.

The example alerts in figure A-4 are typically displayed in a tabular format, figure A-1, and analysts are very good at correlating the data to identify events of interest. In addition to the tabular format, we will also examine at least two more cognitively oriented visual displays, figures A-2 and A-3. During the preliminary and follow-up studies, we will apply full randomization of the test subjects, which amounts to the following possible sequences of the subjects using the visual displays:

- Group 1: A→B→C
- Group 2: A→C→B
- Group 3: B→A→C
- Group 4: B→C→A
- Group 5: C→A→B
- Group 6: C→B→A

Thus, two subjects will perform each ordering of visual displays. While not statistically significant, this should begin to identify any impact of the ordering on performance, which itself may aid in training of future analysts.

The Parallel Coordinates Display was originally generated by a tool called GUESS, an exploratory data analysis and visualization tool for graphs and networks (9). GUESS contains a domain-specific embedded language called Gython, an extension of Jython. Jython is a Java based language derived from Python. Gython supports the operators and syntax necessary for working on graph structures in an intuitive manner. The tool also offers a visualization front end that supports the export of static images and dynamic movies. We selected GUESS because Army Research Laboratory researchers preferred its ease of use; alternative development environments may be used but the displays and interactions will remain essentially similar to what is described. We chose this tool to represent interactions within a network. With the tool, we show a single connection from one system to another as a solid directed line. Large hashes along the directed line represent users who have multiple connections to more than one system. A single hash mark represents each user. With this information, we can measure activity between systems and monitor behavior patterns. We use red to highlight unusual or unexpected activity (11). Figure A-5 for the visual key of these directed line representations.

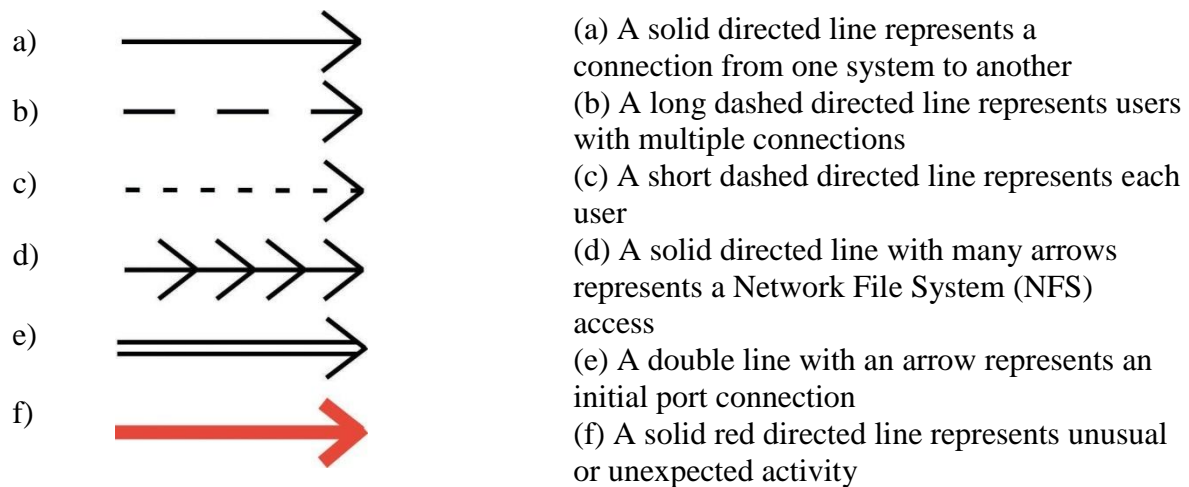


Figure A-5. (a) Line Visualizations Used for Parallel Coordinate Display. (b) Meanings of Line Visualizations Used (11).

For the Node-Link Display, glyph-based visual representations are created as visual attributes to portray connections that could exist within any system. These parameters include but are not limited to number of users, system load, status, and unusual or unexpected activity (11). For the preliminary study, we introduce these glyph representations and ask the participants for their response to effectiveness in communicating network parameters of a system's data. The result is a display of visual attributes that are easily interpretable for their actual meaning. In the follow-up study, the visual attributes are designed in a cyber-defense network game scenario in conjunction with the database parameters in such a way that the correlation is appropriate and the

relationship is comprehensible to the analyst. Figure A-6 for the visual key of the glyphs representations and meaning.

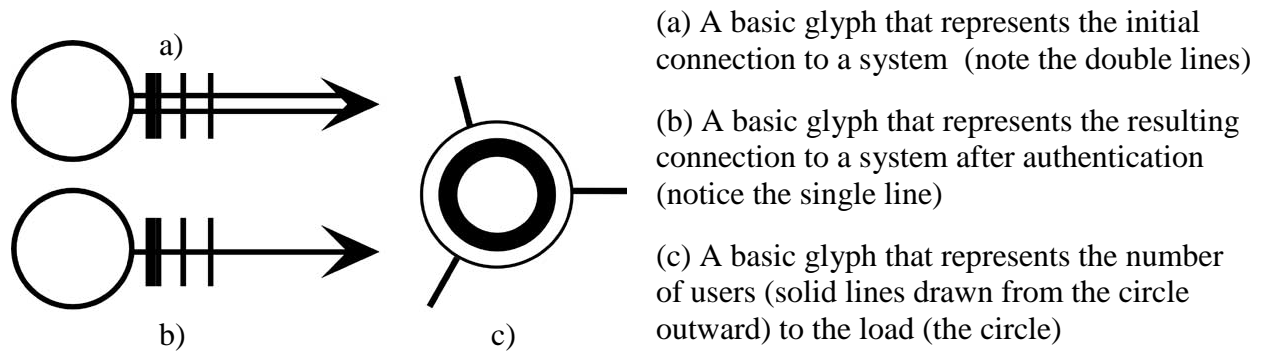
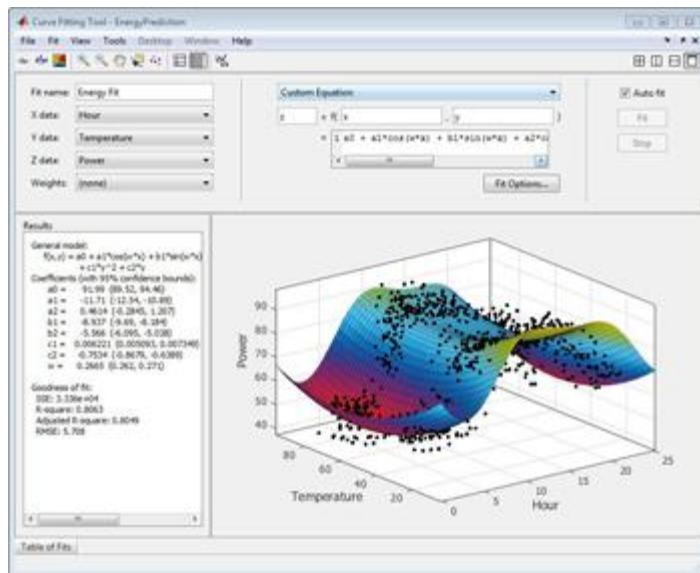


Figure A-6. (a) Line Visualizations Used for Node-Link Display. (b) Meanings of Line Visualizations Used (11).

Preliminary Study (no game scenario)

Participants will detect the intrusions and possible intrusions on the visual representations. Subjects will examine the highlighted features within each visual representation to determine which alerts are intrusions or possible intrusions. They will then provide feedback on the effectiveness of the communication on each visual representation of cyber-defense network data. In this study, an intrusion is defined as the ability to compromise a computer system by breaking the security of the system or by causing it to go into an insecure state. We assess a possible intrusion by the identification of events that occur close together in time.

We use MATLAB as the environment to display the visual representations, see figure A-7.



- MATLAB is good for data analysis and visualization
- Able to acquire data from files, other applications, databases, and external devices
- Able to filter, manage, and preprocess data
- Provides built-in 2D and 3D plotting and volume visualization functions
- Documenting and sharing results are possible via plot or reports
- Reports can be published in a variety of formats, such as HTML, PDF, Word, or LaTeX

Figure A-7. (a) MATLAB shot of plot library being used. (b) MATHLAB highlights.

Follow-Up Study (with the game scenario)

Participants will detect the intrusions and possible intrusions on the simulated network via a game scenario by doing the following:

1. Correlate the different alerts by foreign IP address. A number of different types of alerts coming from the same source are extra suspicious. Try to gather, by sorting, all the traffic to and from the same IP address. A secondary sort should be on the local IP – separate the messaging with different local addresses. The game also has some additional visual graphs and pictures showing traffic volume, separated by foreign IP address.
2. Look for a malware or Trojan name in the “Alert Message Emitted” text. If the alert identifies a particular piece of malware, it is more likely to be a real threat. It would still have to be correlated with other traffic to make it more certain.
3. The foreign country in the source or destination field can usually be calculated based upon the IP address. This is only an indicator, because there is legitimate traffic from unfriendly countries and threatening traffic that appears to come from friendly countries.

We designed a representative dataset that contains a number of ‘alert’ records displayed for an equivalent large site. The fabricated set will contain 500 ‘alert’ records because typically an analyst sees 500 alerts during an hour. However, less than one percent of these alerts actually correlate to an incident or interesting feature of traffic. The game scenario uses the visual displays to illustrate the designed dataset differently. The participants are given instructions on

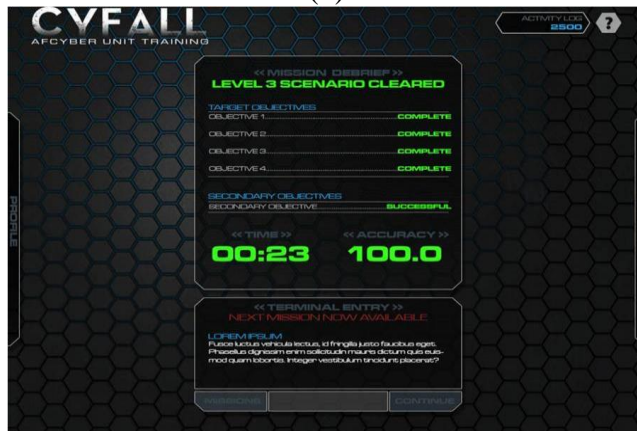
how to play the game, which incorporates the above mentioned tasks. To win the game, a participant must correctly identify all of the intrusions and intrusion attempts on each level (a level represents a visual display). Figure A-8 that shows screenshots of the CyFall tool developed for the game scenario.



(a)



(b)



(c)

Figure A-8. (a) This is a screenshot of the “Introduction Overlay”. It introduces the subject to the mission, provides instructions on identifying threats, and highlights the features/functions of the game. (b) This is a screenshot of the “Exercise Overlay”. Here the subjects will be able to view, explore, and look deeper into the dataset via the particular visualization display. It is here after exploration that the subject determines and identifies the network threat. (c) This is a screenshot of the “Results Overlay”. A running log is kept in the background to keep track of each subject's performance. A module contains all three overlays and repeats three times for the three different visual displays. The subject's performance is displayed at the end of each module and once more at the end of the entire session for their overall time and accuracy performance.

Phase II:

A second goal of this investigation is to begin to understand the process by which analysts perform their analysis. This cognitive process is of great interest since algorithmic approaches have, to date, been unable to duplicate evenly this process remotely. Understanding more about this cognitive process will enable development of tools designed to aid the analysis process as well as the development of algorithms to reproduce said process, particularly in the case of known threats. To this end, a second phase of the protocol will examine an experience-aided reasoning support system developed at Penn State University (PSU) under an Army Research Office (ARO) MURI, see figure A-9. They have designed this tool to both support analysts in the development and evaluation of a hypothesis as well as record the analyst's process of evaluating and rejecting or accepting hypothesis. We will make no association between the recorded processes and the analysts name or identifying characteristics. Again, performance metrics with and without the aid of this tool will be generated for quantitative analysis. In addition to the tool, it is also feasible to have analysts dictate their process, as implied earlier. There is a goal of analyzing the cognitive process tool.

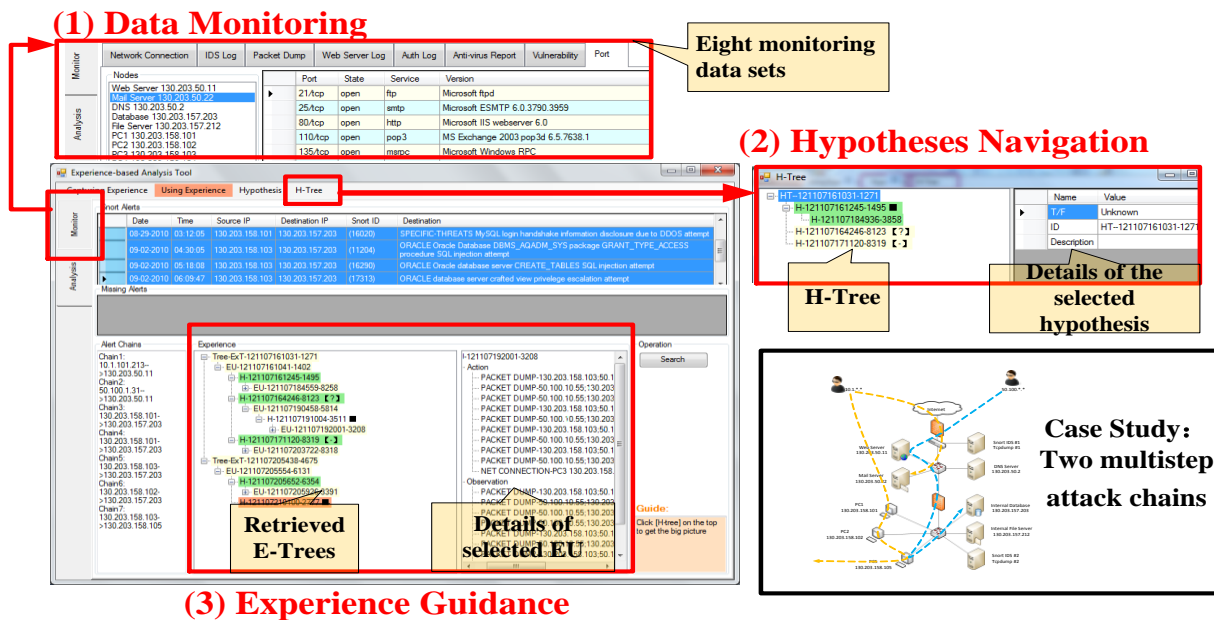


Figure A-9. Experience-aided reasoning support system overview.

Subjects here have a chance to practice analyzing the data provided by the system. We introduce participants to IDS alerts, network configuration, vulnerability reports, data dumps, port scanner reports, and system logs. The subjects are encouraged to speak out aloud during their thinking process of coming up with and finalizing their hypothesis.

Thus for Phase II of the study, we are specifically asking the participants to (for both preliminary and follow-up studies):

1. Conduct analysis on the provided network data.
2. Create hypotheses (notations of thoughts in making decisions within the presented network data).
3. Participants are encouraged to strike original thoughts and make new ones as they would in a natural cognitive thinking process.

The participants' personal performance in this study is not the focus of this research. Instead, their performance helps us to generate better visual representations that maximize saliency of features of interest for network analysts' intrusion detection tasks. In addition, the results aid in building the foundation for the science and theory of intrusion detection. Participants will have a maximum of three hours to complete the tasks in a sitting.

Procedure

1. Preliminary Study

- Step 1. We will begin the study with a welcome followed by an introduction of the investigators.
- Step 2. Investigators will then brief the participants on the study and will obtain informed consent. Participants of this study will be given a random anonymous identification number to protect their personal information and identity. They will be asked to complete a background and demographics questionnaire.
- Step 3. The investigators will explain each visual display used and their specific ways of representing a network system's attributes.
- Step 4. The investigators will ask the subjects to record their thought process in making their decisions. The PSU tool collects these notations as hypotheses (thoughts) and creates a tree that traces a subject's thoughts throughout the experiment.
- Step 5. The investigators will then describe the tools and explain how the participants will use them.
- Step 6. The investigators will conduct a run-through or demo if you will of the participants tasks. This will serve as practice for the participants. We demonstrate how to create a new hypothesis by clicking the mouse on the trace submission button.
- Step 7. The investigators will lead a session to entertain questions that the participants might have concerning their tasks or any other aspects of the study.
- Step 8. Participants will conduct the experiments for Phase I (visual representations only) and Phase II.

- Step 9. Participants will complete their post-task questionnaires and provide the investigators with any final remarks or comments.
- Step 10. Investigators will lead a debrief session and provide the participants with a copy of the signed consent form.

Participants will have a maximum of three hours to complete the tasks in one sitting. See Appendix B for the Pre-Task Questionnaires and Appendix C for General Background Information ask of the participants. See Appendix D through Appendix H for the Post-Task Questionnaires.

2. Follow-Up Study

- Step 1. We will begin the study with a welcome followed by an introduction of the investigators.
- Step 2. Investigators will then brief the participants on the study and will obtain informed consent. Participants of this study will be given a random anonymous identification number to protect their personal information and identity. They will be asked to complete a background and demographics questionnaire.
- Step 3. The investigators will describe the cyber-network game scenario and the participant's associated tasks as a cyber-security analyst.
- Step 4. The investigators will ask the subjects to record their thought process in their making their decisions. The PSU tool collects these notations as hypotheses (thoughts) and creates a tree that traces a subject's thoughts throughout the experiment.
- Step 5. The investigators will then describe the tools and explain how the participants will use them.
- Step 6. The investigators will conduct a run-through or demo if you will of the participants tasks. This will serve as practice for the participants.
- Step 7. The investigators will lead a session to entertain questions that the participants might have concerning their tasks or any other aspects of the study.
- Step 8. Participants will conduct the experiments for Phase I (game scenario) and Phase II.
- Step 9. Participants will complete their post-task questionnaires and provide the investigators with any final remarks or comments.
- Step 10. Investigators will lead a debrief session and provide the participants with a copy of the signed consent form.

Participants will have a maximum of three hours to complete the tasks in one sitting. See appendix B for the Pre-Task Questionnaires and appendix C for General Background Information ask of the participants. See appendix D through appendix H for the Post-Task Questionnaires.

Data Analysis

The performance of the participants will be monitored throughout the entire experiment:

Preliminary Study

We will measure participants' ability to correctly identify suspicious activity (intrusion attempts) and cyber attacks (TP-true positive match for a network threat) using each of the different displays and compare performance among the representations. We also use questionnaires to measure the subject's subjective perception of representations.

Follow-up Study

The time for each participant to complete the entire scenario will be recorded as "Total Time". We will perform statistical analysis by measuring the effectiveness of the comparisons between the input visual displays and by identifying detections versus Total Time. We use log recording to collect the number of intrusion attempts and intrusion attacks correctly identified by the subject. The second metric is the computation of error rate for a strict definition of True Positive as a right answer. The effects of the different visual display types on error rate will be compared. The third metric are scores from the questionnaires themselves to measure the subject's performance. In addition, noted for follow up with the participants, are observable difficulties with the displays or extreme lag time of no action. These results should identify features from the visual graphical displays that are effective for cyber security.

Risks

The study involves minimal risk and minimal discomfort to the participants; the analysts in particular are regularly required to work twelve-hour shifts in front of a computer as part of their assigned duties. The likelihood of any physical, mental, or emotional harm is negligible. There will be no psychologically or physically exhausting work required. The investigators will monitor the safety of the participants in this study however; we cannot eliminate all discomforts that may occur. The following are possible discomforts for this study:

1. Subjects may experience eyestrain in a dimmed light setting during this study.
2. Subjects may experience unexpected discomforts such as sitting discomforts in this study. There is a risk of back pain, leg pain, arm pain, or any other associated pain with sitting for an extended period.

Benefits

There is not an immediate benefit to the participants. However, their contributions and results of the study may improve the quality of the visual display for ARL analysts, allowing them to identify features of interest more efficiently and effectively.

Confidentiality

The participants' personal information remains confidential. The study requires obtaining basic information from participants and no personal information besides a name and signature for the consent form is required. This study uses the participants' responses, performance, and demographic information related to the study in the publication of the research. However, we provide a random anonymous identification number to protect their identity and results for publication. Participants are neither photographed nor videotaped. We will use audio tapes to record their interview responses ensuring clarity and accuracy of their responses. Researchers will review the audio recordings and ensure that no personally identifying information or other sensitive details will be released to the public. Of course, a participant is free to retract their statements during the interview session.

References

1. Visual Analytics Tools for Analysis of Movement Data by G. Andrienko, N. Andrienko, and S. Wrobel. ACM 2007.
2. Visualizing Network Data by R. Becker, S. Eick, and A. Wilks. March 1995.
3. <http://usabilityfriction.com/2010/11/22/cognitive-load/> Usability Friction: Usability shouldn't be a drag Cognitive Load, By Ashley Towers, November 22, 2010 Visited the site on February 7, 2013
4. KMH01- KOSARA R., MIKSCH S., HAUSER H.: Semantic depth of field. *Proceedings of IEEE INFOVIS* (2001), 97–104.
5. TM04- TORY M., MÖLLER T.: Human factors in visualization research. *IEEE Transactions on Visualization and Computer Graphics* 10, 1 (2004), 72–84.
6. Engle, Sophie; Whalen, Sean. Visualizing Distributed Memory Computations with Hive Plots, 56–63. In *Proceedings of the 9th ACM International Symposium on Visualization for Cyber Security*, 2012.
7. Erbacher, Robert F. Visualization Design for Immediate High-Level Situational Assessment. *VizSec '12 Proceedings of the Ninth International Symposium on Visualization for Cyber Security*, pp. 17–24.
8. Anderson, E. W.; Potter, K. C.; Matzen, L. E.; Shepherd, J. F.; Preston, G.; Silva, C. T. A User Study of Visualization Effectiveness Using EEG and Cognitive Load. presented at *Comput. Graph. Forum*, 2011, pp.791–800.
9. Adar, Eytan. GUESS: A Language and Interface for Graph Exploration. *CHI 2006*, Montreal, Canada, Apr. 22–27, pp. 791–800, 2006.
10. Devlin, H. What is Functional Magnetic Resonance Imaging (fMRI)? *Psych Central*. Retrieved on February 26, 2013, from <http://psychcentral.com/lib/2007/what-is-functional-magnetic-resonance-imaging-fmri/>, 2007.
11. Erbacher, R. F.; Walker, K. L.; Frincke, D. A. Intrusion and Misuse Detection in Large-Scale Systems. *Computer Graphics and Applications, IEEE* **2002**, 22 (1), 38–47.
12. Wang, Y.; Luo, L.; Freedman, M. T.; Kung, S. Y. Probabilistic Principal Component Subspaces: A Hierarchical Mixture Model for Data Visualization. *IEEE Transactions on Neural Networks* **2000**, 11 (3), 625–636.
13. Nabney, I. T.; Sun, Y.; Tino, P.; Kaban, A. Semi-Supervised Learning of Hierarchical Latent Trait Models for Data Visualization. *IEEE Transactions on Knowledge and Data Engineering* **2005**, 17 (3), 384–400.

14. Shneiderman, B. Why Not Make Interfaces Better Than 3D Reality? *IEEE Transactions on Computer Graphics and Applications* **2003**, 12–15.
15. Nurse, J. R.; Creese, S.; Goldsmith, M.; Lamberts, K. Trustworthy and Effective Communication of Cybersecurity Risks: A Review. In *Socio-Technical Aspects in Security and Trust (STAST), 2011 1st Workshop on* (pp. 60–68). IEEE, September 2011.
16. Fink, G. A.; North, C. L.; Endert, A.; Rose, S. Visualizing Cyber Security: Usable Workspaces. In *Visualization for Cyber Security, 2009. VizSec 2009. 6th International Workshop on* (pp. 45–56). IEEE, October 2009).
17. http://www.informationbuilders.com/new/newsletter/9-2/05_lozovsky
18. Rasmussen, J.; Ehrlich, K.; Ross, S.; Kirk, S.; Gruen, D.; Patterson, J. Nimble Cybersecurity Incident Management Through Visualization and Defensible Recommendations. In *Proceedings of the Seventh International Symposium on Visualization for Cyber Security* (pp. 102–113). ACM, September 2010.
19. Goodall, J. R. Visualization is Better! A Comparative Evaluation. *Proceedings of the Workshop on Visualization for Computer Security (VizSec)*, IEEE Press, 2009, 57-68, 2009.
20. Koike, H.; Ohno, K. SnortView: Visualization System of Snort Logs. In *Proceedings of the 2004 ACM workshop on Visualization and data mining for computer security* (pp. 143-147). ACM, October 2004.
21. Komlodi, A.; Goodall, J. R.; Lutters, W. G. An Information Visualization Framework for Intrusion Detection. In *CHI '04 Extended Abstracts on Human Factors in Computing Systems* (Vienna, Austria, April 24-29, 2004). CHI '04. ACM, New York, NY, 1743. DOI=<http://doi.acm.org/10.1145/985921.1062935>, 2004.
22. <http://www.mathworks.com/products/matlab/>, accessed on August 22, 2013.

Appendix A: Consent Form

Site of Research: Building 459, Room 202 System Assessment and Usability Laboratory (SAUL), Aberdeen Proving Ground, MD

RESEARCH PARTICIPANT CONSENT FORM ARMY RESEARCH LABORATORY

Project Title: Evaluation of the Presentation of Network Data via Visualization Tools for Network Analysts

Sponsor: Department of Defense

Co-Principal Investigator: Renée E. Etoty, Adelphi Laboratory Center, MD, 301-394-1835, renee.e.etoty.civ@mail.mil

Co-Principal Investigator: Robert F. Erbacher, Adelphi Laboratory Center, MD, 301-394-1674, robert.f.erbacher.civ@mail.mil

Associate Investigator: Christopher Garneau, Aberdeen Proving Ground, MD, 410-278-5814, christopher.j.garneau.civ@mail.mil

Date: 03 April 2013

We are asking you to join a research study. This consent form explains the research study and your part in it. Please read this form carefully before you decide to take part. You can take as much time as you need. Please ask questions at any time about anything you do not understand. You are a volunteer. If you join the study, you can change your mind later. You can decide not to take part right now or you can quit at any time later on.

Why is this research being done?

We invite you to participate in a study designed to assess visual layouts for cyber-security network analysts on representative network activity; in essence, we will use computer graphical displays to represent computer network activity that the network analysts currently view in a tabular format. This study will examine the cognitive aspects of visual displays with the goal of identifying representations and components of representations that most effectively aid network analyst in interpreting the underlying activity in a network data sample. The Army Research Laboratory (ARL) – Computational Sciences, Information Directorate (CISD), and Human Research Engineering Directorate (HRED), are conducting the study.

What will happen if you join this study?

As a participant of this study, we will give you a random anonymous identification number to protect your personal information and identity. We will ask you to complete a background experience and demographics questionnaire. An investigator will describe the tasks for the Preliminary Study or Follow-Up Study. We describe the specific tools and tasks below. You will

have a maximum of three hours to complete the tasks in a sitting. There will be a fifteen-minute session after a short demonstration-training period to address your questions, comments, and concerns. Upon completion of the tasks, we will ask you to execute a post-task questionnaire that will reflect your comments about the overall study experience and your impression of the visual displays used.

Particularly for the Preliminary Study, specifically, we ask you to detect the intrusions and possible intrusions on the visual representations provided. You will examine the highlighted features within each visual representation to determine which alerts are intrusions or intrusion attempts. You will then provide feedback on the effectiveness of the communication on each visual representation of cyber-defense network data. In this study, an intrusion is defined as the ability to compromise a computer system by breaking the security of the system or by causing it to go into an insecure state. We assess a possible intrusion by the identification of events that occur close together in time.

For the Follow-Up Study, we ask you to play a cyber-network game scenario that incorporates a pattern matching behavior representative of a typical cyber-security analysis session. Your goal for the game is to identify all the intrusions and intrusion attempts made by the cyber attacker. This technique is coupled with a visual display that aids an analyst in performing their tasks. The visual task scenarios of the game will compare analyst effectiveness across at least three constructed visual layouts.

We provide you with the tools used to carry out your tasks for both studies. These tools consist of at least three types of displays showing network activity of interest to network analysts. The first display is a sort-able table. The second display is a colored parallel coordinate representation of alerts and normal traffic with a data inspector pane. The third display is a “node-edge” representation. A fabricated set of data that contains 500 ‘alert’ records will be presented in the study an analyst typically sees about 500 alerts during an hour. You will be able to manipulate this dataset on each visual display to help better identify intrusions on the simulated network in the game.

The second phase of the experiment of both studies requires you to practice developing a hypothesis (theory) of the status of the given system. We give you the opportunity to conduct an evaluation of your developed hypothesis. We will record your process of evaluating and rejecting or accepting the hypothesis using log files.

Your personal performance in this study is not the focus of this research. Instead, your performance helps us to generate better visual representations that maximize saliency of features of interest for network analysts’ intrusion detection tasks. Also, note that we will record your performance and we will in no way disclose this information to your respective communities nor publish any identifying information that is traceable back to you.

How much time will the study take?

Your participation in this study will take up to a maximum of three hours for one sitting.

What are the risks or discomforts of the study?

The likelihood of any physical, mental, or emotional harm is remote. There will be no psychologically or physically exhausting work required. The investigators will monitor your safety however, we cannot eliminate all discomforts that may occur. The following are possible discomforts for this study:

1. You may experience eyestrain in a dimmed light setting during this study.
2. You may experience discomforts due to sitting for an extended period during this study. (e.g., there is a risk of back pain, leg pain, arm pain, or any other associated pain with sitting for an extended period).

Are there benefits to being in the study?

To you as the participant, there is no immediate benefit for participating in this study. Your participation as a student subject or an expert subject allows us to use your results and feedback to improve the generation of visual representations that maximize saliency of features of interest for network analysts. This leads to better quality of the visual displays for cyber-security analysts, allowing them to identify features of interest more efficiently and effectively.

Will you be paid if you join this study?

You will receive no payment for taking part in this study.

How will your privacy be protected?

We will keep your personal information confidential. Your personal information will be stored and secured in a locked and password protected computer at our study site. After transfer of your personal information to our secured computer, we will shred the paper copies containing your personal information. In addition, we provide a random anonymous identification number to protect your identity and your results. Publication of the results of this study in a journal or technical report or presentation at a meeting will not reveal personally identifiable information. We will neither photograph nor videotape you. The investigators will further protect your personal information from disclosure to individuals not connected with this study. However, we cannot guarantee complete confidentiality because law permits officials of the U. S. Army Human Research Protections Office and the Army Research Laboratory's Institutional Review Board to inspect the records obtained in this study to insure compliance with laws and regulations covering experiments using human subjects. The principal investigator will retain this consent form for a minimum of three years.

Indicate below if we have your permission to audio record you during the experimental session. We will use audio recordings to record your interview responses ensuring clarity and accuracy of your responses. Please indicate below if you will agree to allow us to record you. You can still participate in this study if you prefer not to be audio recorded.

I give consent to be audio taped during this study: ___Yes ___No please initial: ____

Where can I get more information?

You have the right to obtain answers to any questions you might have about this research both while you take part in the study and after you leave the research site. Please contact anyone listed at the top of the first page of this consent form for more information about this study. You may also contact the chairperson of the Human Research & Engineering Directorate, Institution Review Board, at (410) 278-5992 with questions, complaints, or concerns about this research, or if you feel this study has harmed you. The chairperson can also answer questions about your rights as a research participant. You may also call the chairperson's number if you cannot reach the research team or wish to talk to someone who is not a member of the research team.

Voluntary Participation

Your decision to be in this research is voluntary. You can stop at any time. You do not have to answer any questions you do not want to answer. Refusal to take part in or withdrawal from this study will involve no penalty or loss of benefits you would receive by staying in it.

Military personnel cannot be punished under the Uniform Code of Military Justice for choosing not to take part in or withdrawing from this study, and cannot receive administrative sanctions for choosing not to participate. Civilian or contractor personnel cannot receive administrative sanctions for choosing not to participate in or withdrawing from this study. Once we have answered your questions about the study, and if you want to continue your participation in this study, please sign below.

WE WILL GIVE YOU A COPY OF THIS CONSENT FORM

Signature of Participant	Printed Name	Date
--------------------------	--------------	------

Signature of Person Obtaining Consent	Printed Name	Date
---------------------------------------	--------------	------

Appendix B: Pre-Task Questionnaire

Demographic Information

- 1) What is your gender?
 - a. Male
 - b. Female

- 2) What is your race?
 - a. American Indian or Alaska Native
 - b. Asian
 - c. Black or African American
 - d. Native Hawaiian or Other Pacific Islander
 - e. Other
 - f. White
 - g. Prefer not to say

- 3) What is your age?
 - a. 18-25 years old
 - b. 26-35 years old
 - c. 36-45 years old
 - d. 46-55 years old
 - e. 56-65years old
 - f. 66-75years old
 - g. 76 years or older

- 4) What is the highest level of education you have completed?
 - a. Elementary school only
 - b. Some high school, but did not finish
 - c. Completed high school
 - d. Some college, but did not finish
 - e. Two-year college degree / A.A / A.S.
 - f. Four-year college degree / B.A. / B.S.
 - g. Some graduate work
 - h. Completed Masters or professional degree
 - i. Advanced Graduate work /Ph.D.

- 5) What is your work title?

- 6) What is your current department?

7) Do you have any vision impairments or poor vision (after correction)?

a. Yes

b. No

c. If yes, please explain.

8) Do you have any other disabilities?

a. Yes

b. No

c. If yes, please explain.

Appendix C: General Background Information

- 1) Do you use computers (PC's, MAC, iPad, iPhone, Android phone, tablets, etc.)?
 - a. Yes
 - b. No
- 2) How often on a daily basis, do you use computers?
 - a. 1-5 hrs
 - b. 5-10 hrs
 - c. 10-15 hrs
 - d. 15-20 hrs
 - e. More than 20 hrs
- 3) How comfortable do you feel using a computer?
 - a. Very comfortable
 - b. Somewhat comfortable
 - c. Somewhat uncomfortable
 - d. Very uncomfortable
- 4) Have you ever written a software program or a mini computer code?
 - a. Yes
 - b. No
- 5) Have you ever configured a Linux computer?
 - a. Yes
 - b. No
- 6) What is a shell?

- 7) Have you ever worked as a network analyst or have any network analysis experience? If so, state where and when.
 - a. Yes
 - b. No
 - c. State your experience.

If an analyst, answer questions #8 and #9.

- 8) How many years have you been a cyber analyst?
- a. Less than 1 year
 - b. 1 to 3 years
 - c. 3 to 5 years
 - d. 5 to 10 years
 - e. More than 10 years
 - g. Never
- 9) Which of the following activities do you most frequently perform in your work?
Check all that apply.
- ☐ Filter raw sensor data (e.g. IDS alerts).
 - ☐ Point out the suspicious activities from filtered data.
 - ☐ Collect evidence from multiple sources (e.g. IDS, package dumps, etc.).
 - ☐ Group individual activities and make hypotheses about an intrusion attempt.
 - ☐ Assess attacker identity and mission impact
 - ☐ Tuning sensors to look for predicted attack
 - ☐ Incident handling
 - ☐ Produce documents to report current situation awareness
 - ☐ Perform virus/incident handling
 - ☐ Train others of situation awareness

Appendix D: Post-Task Questionnaire

Subjective Survey # 1

For the following questions, indicate with your opinion on a scale of 1 through 5.

1 – Strongly Disagree
2—Somewhat Disagree

3 – Neutral

4 – Somewhat Agree
5 – Strongly Agree

- 1) The graphical displays were visually more appealing to the eye than the tabular display. _____
- 2) I easily understood the visualization of the graphical displays. _____
- 3) I easily understood the visualization of the tabular display. _____
- 4) The manipulation of the visualization's features of the graphical displays was easy. _____
- 5) The manipulation of the visualization's features of the tabular display was easy. _____
- 6) I was able to identify all of the *intrusion* alerts on the graphical displays. _____
- 7) I was able to identify all of the *intrusion* alerts on the tabular displays. _____
- 8) I was able to identify all of the network *intrusions* on the graphical displays. _____
- 9) I was able to identify all of the network *intrusions* on the tabular display. _____
- 10) The demo training provided by the investigators enabled me to use effectively the tool. _____
- 11) I was able to complete my tasks better with the tabular display than the graphical displays. _____
- 12) I prefer the graphical displays to the tabular display. _____
- 13) I recommend that the use of the *graphical displays* along with the GUESS visualization tool be incorporated into analyst's cyber-security systems. _____
- 14) I recommend that the use of the *tabular display* along with the GUESS visualization tool be incorporated into analyst's cyber-security systems. _____
- 15) I do not recommend that the *graphical displays* along with the use of the GUESS visualization tool be incorporated into analyst's cyber-security systems _____

- 16) I do not recommend that the *tabular display* along with the use of the GUESS visualization tool be incorporated into analyst's cyber-security systems. _____
- 17) The phase two system provides me helpful guidance by suggesting relevant experience pieces. _____
- 18) Most experience pieces suggested by the phase two system are relevant. _____
- 19) The representation of experience in the phase two system is easy to understand. _____
- 20) Interacting with the phase two system distracts me from concentrating on reasoning. _____
- 21) The display of the phase two system helped me manage my hypotheses and was very useful. _____
- 22) Generally, the phase two system makes a positive impact on my reasoning process. _____

Appendix E: Post-Task Questionnaire

Subjective Survey #2

1. Overall, how would you rate the usefulness of the graphical layouts?

- ☐ Excellent
- ☐ Good
- ☐ Fair
- ☐ Poor

2. Overall, how would you rate the appearance of the tabular layout?

- ☐ Excellent
- ☐ Good
- ☐ Fair
- ☐ Poor

3. What components of the displays were most effective?

4. What aspects of the visualizations did you like best?

5. What aspects of the visualization did you not like?

6. What aspects of the visualization s helped you to identify intrusions?

7. Do you think the phase two system can really help analysts do analytical reasoning in cyber analysis tasks? What advantages does it have?

8. What three things did you like **most** about your interactions with the phase two system today?

9. What three things did you like **least** about your interactions with the phase two system today?

Appendix F: Post-Task Questionnaire

Analysis Survey #1

Answer the following questions based on the node representations provided in FIGURE 1 below.
Indicate your selections by writing the letter of the node representations in the blank line.

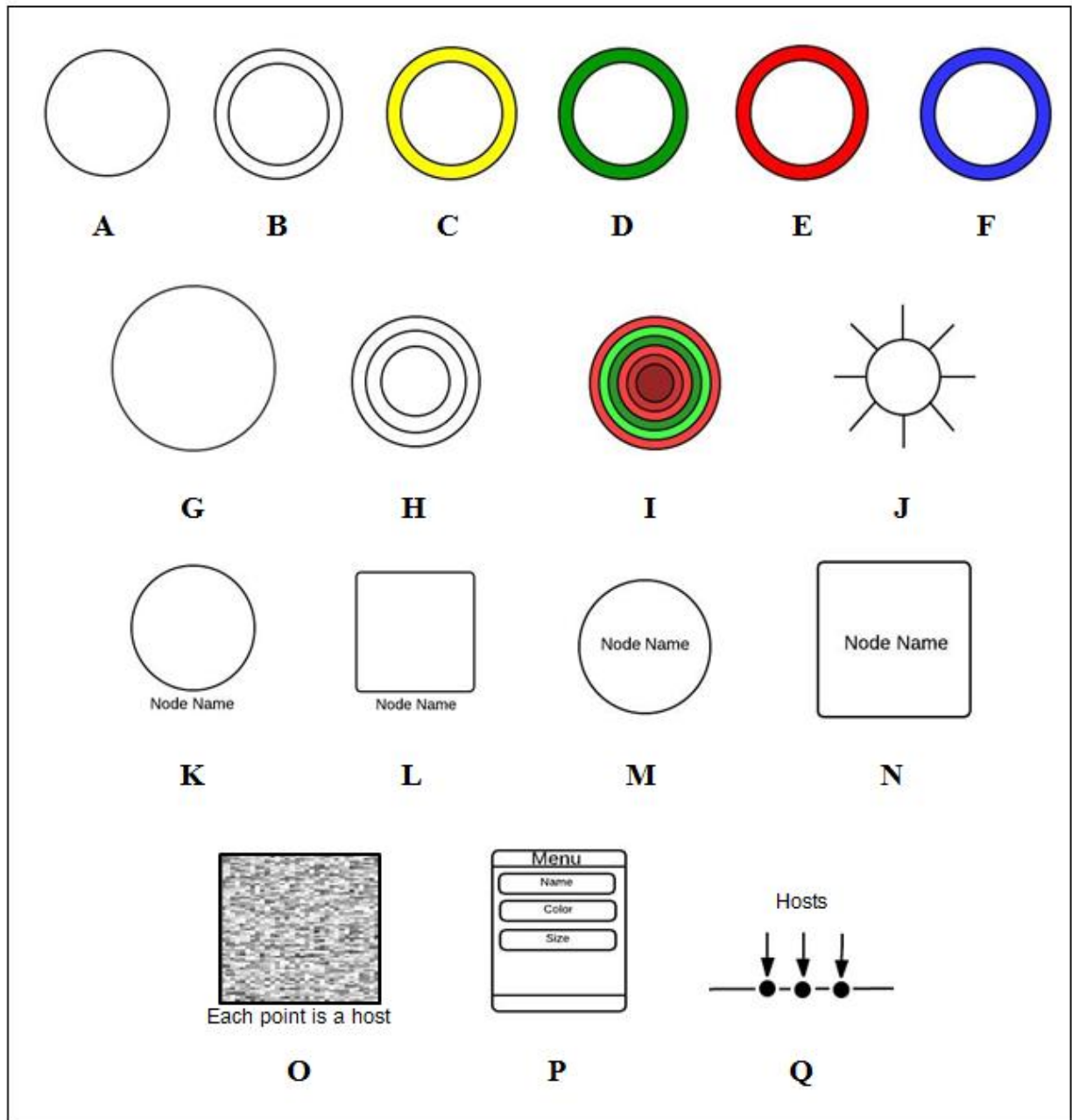


FIGURE1: Node Representations

Note(The same representation can be used in multiple answers.)*

- 1) From the set (G, H, I, J, O, P), which representation is best suited for representing the activity of the system, i.e., top talker? _____
- 2) From the set (A, K, L, M, N, P), which representation is best suited for labeling a system? _____
- 3) From the set (C, D, E, F, G, I, J, O, P), which representation is best suited for representing the number of users located at a system? _____
- 4) From the set (H, I, J, P), which representation is best suited for representing the relevant past history of a system? _____
- 5) Which representation is best suited to represent an active system? _____
- 6) Which representation is best suited to represent an inactive system? _____
- 7) Which representation is best suited to represent a system under attack? _____
- 8) Which representation is best suited to represent a system that is vulnerable? _____
- 9) Which representation is best suited to represent a system that has been compromised? _____
- 10) Which representation is best suited to represent a high priority system? _____
- 11) Which representation is best suited to represent a low priority system? _____
- 12) Prioritize the following network parameters in terms of relevance to analysis, 1 being highest priority, 13 being lowest priority, NA means it is not used/relevant/of interest.
 - _____ CPU Load
 - _____ Number users
 - _____ Number connections
 - _____ Network bandwidth usage
 - _____ % Disk usage
 - _____ % memory usage
 - _____ # alerts generated
 - _____ Type of alerts generated
 - _____ Previous identification of issues with a specific system
 - _____ Median size of packets
 - _____ Connections asymmetry
 - _____ Operating system type
 - _____ System priority
 - _____ Other(s): _____

Appendix G: Post-Task Questionnaire

Analysis Survey #2

Answer the following questions based on the link representations provided in FIGURE 1 below.
Indicate your selections by writing the letter of the link representations in the blank line.

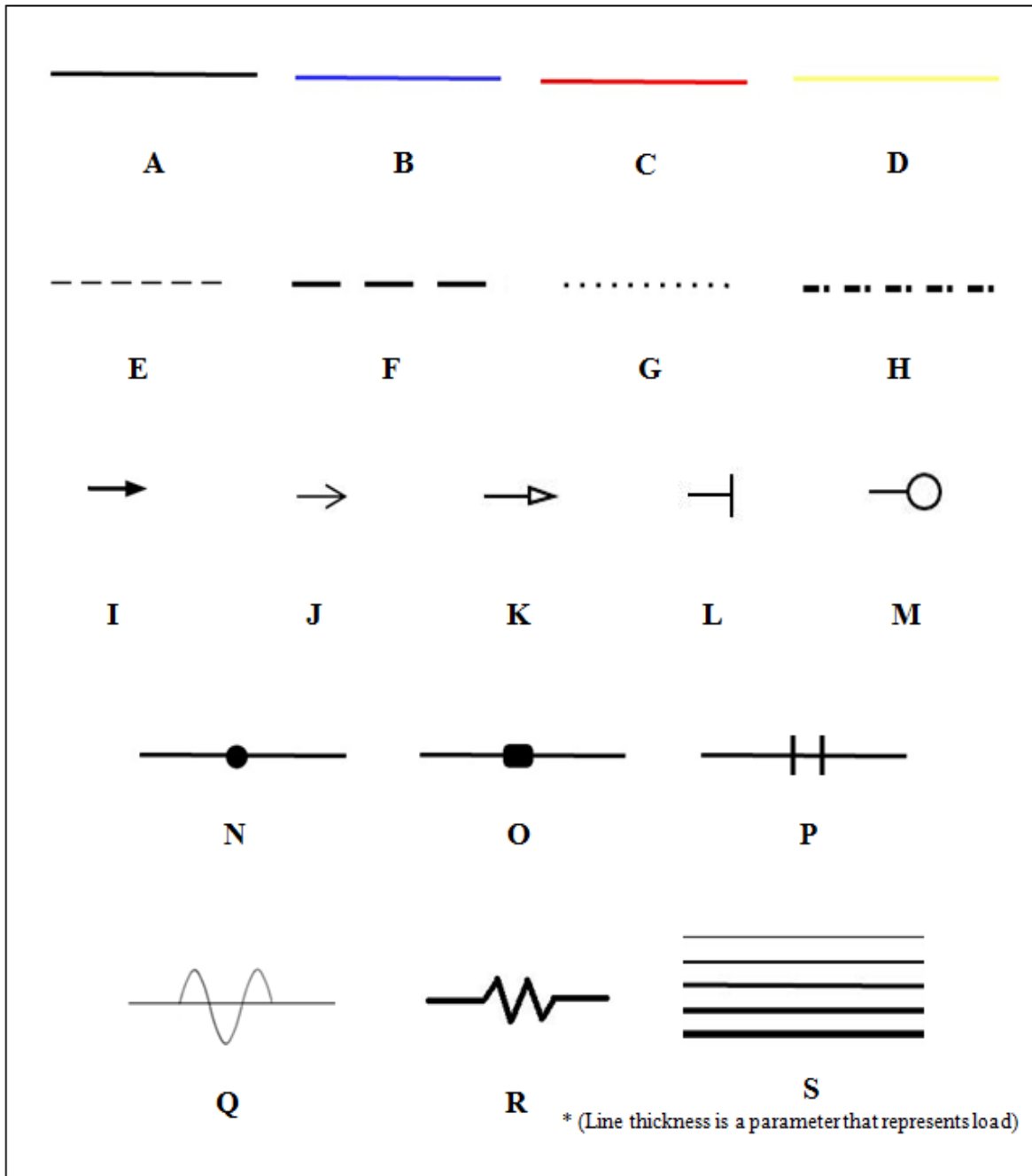


FIGURE1: Link Representations

Note(The same representation can be used in multiple answers.)*

- 1) Which representation is best suited to represent a connection from one system to another system? _____
- 2) Which representation is best suited for representing users with multiple connections?

- 3) Which representation is best suited to represent a TCP connection? _____
- 4) Which representation is best suited to represent a UDP connection? _____
- 5) Which representation is best suited to represent access to a Network File System (NFS)?

- 6) From the set (I...M), which representation is best suited for representing connections to a server? _____
- 7) From the set (I...M), which representation is best suited for representing connections to a client? _____
- 8) From the set (I...M), which representation is best suited for representing connections to a UNIX system? _____
- 9) From the set (I...M), which representation is best suited for representing connections to a Windows system? _____
- 10) Which representation is best suited for representing CONUS connections? _____
- 11) Which representation is best suited for representing OCONUS connections? _____
- 12) Which representation is best suited to represent activity that generated an alert? _____
- 13) Which representation is best suited to represent the connection **from** a system under attack? _____
- 14) Which representation is best suited to represent the connection **to** a system under attack?

- 15) Which representation is best suited to represent an unauthorized system connection?

- 16) Which representation is best suited to represent normal traffic communications between systems? _____
- 17) Which representation is best suited for asymmetry of connections between inbound and outbound? _____
- 18) Which representation is best suited for representing the number of connections over the past 5 minutes? _____
- 19) Which representation is best suited for representing the number of connections over the past 1 hour? _____
- 20) Which representation is best suited for representing the number of connections over the past 24 hours? _____

Appendix H: Pre-Task Questionnaire

Analysis Survey #3

(Before execution of the study)

Answer the following questions to the best of your abilities.

- 1) On a scale from 1 to 5 where 1 is 'Very Sad', 3 is 'Neutral', and 5 is 'Very Happy', which of the following best describes your current emotional state?
 - ☐ 1-Very Sad
 - ☐ 2-Sad
 - ☐ 3-Neutral
 - ☐ 4-Happy
 - ☐ 5-Very Happy

- 2) Which is better at the following tasks, Machine or Human? Write your answer in the blank line.
 - ☐ Analyzing data _____
 - ☐ Detecting anomalies (where an anomaly is an abnormal behavior on a security network) _____

- 3) In general, which type of display do you find most useful in analyzing data? Check all that apply.
 - ☐ Tables
 - ☐ Textual
 - ☐ Line & Bar Graphs
 - ☐ Simple Graphs
 - ☐ Symbolic shapes
 - ☐ Other: _____

- 4) If you are not a cyber security analyst, are you interested in what they do?
 - ☐ Highly interested
 - ☐ Somewhat interested
 - ☐ Unsure
 - ☐ Not quite interested
 - ☐ Not at all interested

- 5) What motivated you to participate in this study?

6) What do you expect to learn from this study?

Appendix I: Post-Task Questionnaire

Analysis Survey #4

(After execution of the study)

Answer the following questions to the best of your abilities.

- 1) On a scale from 1 to 5 where 1 is 'Very Sad', 3 is 'Neutral', and 5 is 'Very Happy', which of the following best describes your current emotional state?
 - ☐ 1-Very Sad
 - ☐ 2-Sad
 - ☐ 3-Neutral
 - ☐ 4-Happy
 - ☐ 5-Very Happy
- 2) Which is better at the following tasks, Machine or Human? Write your answer in the blank line.
 - ☐ Analyzing data _____
 - ☐ Detecting anomalies (where an anomaly is an abnormal behavior on a security network) _____
- 3) Which types of displays do you find most useful in analyzing security data? Check all that apply.
 - ☐ Tables
 - ☐ Textual
 - ☐ Line & Bar Graphs
 - ☐ Simple Graphs
 - ☐ Symbolic shapes
 - ☐ Other: _____
- 4) If you are not a cyber security analyst, how likely now are you to become one?
 - ☐ Very Likely
 - ☐ Likely
 - ☐ Unsure
 - ☐ Not Likely
 - ☐ Very Unlikely
- 5) If you are a cyber security analyst, how likely are you to be one in the future?
 - ☐ Very Likely
 - ☐ Likely
 - ☐ Unsure
 - ☐ Not Likely
 - ☐ Very Unlikely
- 6) What features of the visual displays were most useful for completing your tasks in this study?

- 7) What did you learn from this study?

Appendix J: Email recruitment letter for ARL analysts

Good morning/afternoon,

The U.S. Army Research Laboratory (ARL) Computational and Information Sciences Directorate (CISD) and Human Research and Engineering Directorate (HRED) are seeking adults (ages 18 and above) with cyber security network analysis experience to participate in a research study evaluating effective visual displays for analysts. In the research, we employ computer graphical displays to represent computer network activity that network analysts currently view in a tabular format. We are interested in participants' response to a simulated cyber-security analysis game scenario. During the study, participants act as analysts and their job is to identify as many of the intrusion attacks and intrusion attack attempts on a simulated network using tabular and graphical displays. We will use the results from the study to help understand which of the visual layouts is most effective for data analysis prediction. This new insight is beneficial for network security analysts tasked with defending the nation's networks from cyber attacks.

If you elect to take part in the research study and are an employee of ARL, you will participate during your regular tour of duty for a maximum of 1 hour per day during a maximum of 5 days. We expect the study to take 2-3 hours for most participants, depending on how quickly tasks are completed and how many rest breaks are taken. There is no compensation or personal benefit for your participation in this study. The study will take place on Aberdeen Proving Ground (APG) in Building 459. Transportation will be provided from other locations at APG, and gate access will be coordinated prior to the study. You can withdraw from this study at any time. Even if you come to the research site and start the study, you can change your mind and withdraw from the study without penalty.

If you would like additional information, please contact the principal investigators:

Renée Etoty

Network Security Branch, ARL Computational and Information Sciences Directorate
Building 204, Room 2D068, Adelphi Laboratory Center, MD
(301) 394-1835
renee.e.etoty.civ@mail.mil

Dr. Robert Erbacher

Network Security Branch, ARL Computational and Information Sciences Directorate
Building 204, Room 2C100, Adelphi Laboratory Center, MD
(301) 394-1674
robert.f.erbacher.civ@mail.mil

Project Number: ARL 13-050

<u>NO. OF COPIES</u>	<u>ORGANIZATION</u>
1 (PDF)	DEFENSE TECHNICAL INFORMATION CTR DTIC OCA
2 (PDFS)	DIRECTOR US ARMY RSRCH LAB RDRL CIO LL IMAL HRA RECORDS MGMT
1 (PDF)	GOVT PRINTG OFC A MALHOTRA
3 (PDFS)	DIR USRL RDRL CIN D R E ETOTY R F ERBACHER DR. CHRISTOPHER GARNEAU